Identifying Glaucoma Susceptibility Genes in Extended Families

by

Patricia Stacey Graham

BSc(Hons), GradDipEd

Menzies Institute for Medical Research | College of Health and Medicine

Submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy (Medical Studies)

University of Tasmania, March 2020

Declaration of Originality

This thesis contains no material which has been accepted for a degree or diploma by the University or any other institution, except by way of background information and duly acknowledged in the thesis, and to the best of my knowledge and belief no material previously published or written by another person except where due acknowledgement is made in the text of the thesis, nor does the thesis contain any material that infringes copyright.

13/3/2020

Authority of access

13/3/2020

Statement of ethical conduct

The research associated with this thesis abides by the international and Australian codes on human experimentation and the guidelines and the rulings of the Ethics Committee of the University.

Ethics Approval Numbers:

University of Tasmania Human Research Ethics Committee H0014085

Oregon Health and Science University Institutional Review Board #1306

13/3/2020

Dedication

To my parents

Susie and Micky Lowry

Who instilled in me the love of learning

and who would have been so proud

Acknowledgements

What a rollercoaster the last four years have been …. it's been an amazing ride. Rollercoasters are no fun alone, and I would sincerely like to thank the following people who kept the ride rolling and to those who sat with me in the carriage and didn't let me fall out.

To the Australian Government for my Research Training Program Scholarship and to the Menzies Institute for supporting me and providing the opportunity to get back into medical research again.

To my supervisors Professor Kathryn Burdon and Dr Jac Charlesworth; the dynamic duo of supervisory teams. Kath, you are a "mentor extraordinaire". Your intelligence, passion and kindness are the perfect combination. You've always had the confidence in me that sometimes I didn't have in myself, and I truly appreciate that. Jac, thank you for trusting me with your POAG families and teaching me linkage. You are one of the most patient people I know and you have taught me so much. You've constantly challenged me, which has made me a better scientist. Thank you both for taking me on, teaching me so much and getting me to the end.

To the participants of OGGS and GIST, without whom this study would not have been possible.

To the researchers, clinicians and statisticians involved with this project and to those who kindly provided us with data: Professors John Blangero, Jamie Craig, Joanne Curran, Alex Hewitt, Stuart MacGregor, David Mackey, Mary Wirtz, and Dr Tiger Zhou.

To Sionne and JoJo – my amigos. It's been the best ride with you the past four years. You have both been super patient with my "old school genetics" and getting me up to speed. More importantly, I've loved our morning chats, procrasti-coffees, lunches in the roundabout, time spent at various sports and having the odd celebratory drink. Cheers!

To the team of bioinformaticians who have taught me so much along the way. Juan, Nick and Bennet, I absolutely could not have done this project without you. An extra thank you to Juan for helping me troubleshoot my methods and get them up and running.

To the CompGen team and my "office-kids" in 502 – Ming, Duran, Raj, Kelsie, Alex, Aparna, Emma and Van. You've kept me young!

To my friends from outside Menzies. Although many of you never really understood why I would do this, you've been nothing but supportive. Special thanks to my coffee-buddy, Karen.

Most importantly, to my family; Al, Jess and Nicki. I could not have done this without your support, encouragement and love. I promise – no more mid-life crises …. for a while at least.

Abstract

Glaucoma encompasses a heterogenous group of eye diseases and is the leading cause of irreversible blindness worldwide. Death of retinal ganglion cells causes degeneration of the optic nerve with resultant loss in vision. Primary open-angle glaucoma (POAG) is the most common form of glaucoma. An increase in pressure inside the eye, the intraocular pressure (IOP), is the main risk factor for developing POAG. Pressure reducing medication and surgery are currently the only preventative measures, and means of slowing the progression of the disease, available to those at risk of POAG or already diagnosed with this disease. There is currently no cure for glaucoma.

POAG is a highly heritable, complex disease. Genome-wide association studies (GWAS) have identified many common risk variants, each with small effect sizes, which may increase an individual's susceptibility to developing this disease. However, these variants together only account for around 3% of the heritability of POAG. Rare variants have been proposed as a source of the missing heritability, and the most powerful way to enrich for rare variants is through family studies. Linkage studies have identified genes harbouring rare variants which cause the disease in around 10% of cases. It is proposed that other, unidentified rare variants with large effect sizes, may be important in POAG pathogenesis.

The clinical intermediate traits of POAG that have been used for successful genetic studies include IOP, vertical cup to disc ratio (VCDR) and central corneal thickness (CCT). These highly heritable, quantitative traits are measurable in all individuals, regardless of their POAG disease status. IOP and VCDR are significantly genetically correlated with POAG disease status, making them ideal endophenotypes to use to identify variants and genes involved with POAG pathogenesis. CCT is not significantly genetically correlated with POAG, but is a well-recognised risk factor for the disease, with thinner corneas associated with an increased risk of developing POAG. Identifying genes important in corneal development still gives us the opportunity to better understand the risk of developing POAG. GWAS have been very successful in identifying variants associated with each of these clinical intermediate traits, however, only a small proportion of the overall heritability for each trait is accounted for by these variants.

For this study, we hypothesised that rare genetic variants associated with heritable ocular clinical traits could be identified in extended pedigrees and would improve our understanding

of the genetic susceptibility to POAG. The first aim of this study was to determine the contribution of published genetic loci associated with POAG and related traits, to the trait variance of the clinical phenotypes in five extended POAG enriched families. The next aim was to use linkage analysis to identify regions harbouring potential POAG related risk variants in these five families. The final aim was to use *in silico* tools to investigate genetic variants within the most significant peaks identified from the linkage analysis. This study utilised whole exome sequencing data from 249 members of five large POAG enriched families, with detailed clinical data including IOP, VCDR and CCT measurements, to perform family-based association and linkage analyses within a variance components framework. These five extended families provide the opportunity to identify rare variants involved in disease due to the potential for variants to be enriched by transmission from founder individuals to successive generations.

To identify the previously reported POAG loci, Aim 1 of this study began with a comprehensive literature review. To determine whether these loci contributed to the variance of the clinical intermediate traits of the families in this study, a family-based association analysis was conducted on genetic variants within literature reported regions. Several variants within published linkage regions were significantly associated with each of the three traits; including a single rare variant within *MST1* on chromosome 3 ($p = 9.60x10^{-5}$) and a haplotype within *FLRT3* on chromosome 20 ($p = 1.17x10^{-4}$ to $p = 3.18x10^{-4}$), which were associated with a large increase in IOP in the family members who carry them. *FLRT3* lies in a previously mapped glaucoma linkage region, GLC1K, and is an ideal candidate gene for further study.

To identify novel genes associated with POAG and its endophenotypes, Aim 2 focused on linkage analyses using whole exome sequencing data, which were conducted for the IOP and VCDR traits in the five families. These are both significantly heritable traits, with heritabilities calculated at 52% ($p = 5.13x10^{-10}$) for IOP and 41% ($p = 9.0x10^{-7}$) for VCDR in the families of this study. Identity by descent relationships between family members were calculated directly from the whole exome sequencing data using IBDLD, a program specifically designed for dense genotyping data. Variance components linkage analysis was conducted using SOLAR, including correction for both the ascertainment of the POAG enriched families used in this study, and non-normal trait distributions. Linkage peaks reaching significance were identified for IOP, but not VCDR. For IOP, peaks on chromosomes 9q34.3 and 15q11.2-13.2 reached full significance (maximum LOD = 4.25 and 3.40 respectively), and peaks on

chromosomes 2q23.1-31.1, 3p13-12.1, 6q24.2-25.1 and 7p11.2-q11.23 reached suggestive significance (maximum LOD > 1.86). All but one of the peaks were novel, with the chromosome 15 peak completely overlapping the previously identified GLC1I locus.

Aim 3 involved analysing the identified linkage peaks using *in silico* strategies. Variants within each region were prioritised based on their family-based association results, their effect on reducing the linkage peak when included in the linkage model, and predicted *in silico* functional effects. Candidate genes proposed by these analyses were further investigated by a pathways analysis and were assessed for association in a large GWAS for IOP. *PKP4* and *FMNL2*, on chromosome 2, and *HERC2*, on chromosome 15, are proposed as promising candidate genes associated with increased IOP in the families of this study. *FMNL2* has already been associated with increased IOP and with POAG in recent GWAS. The identification of *FMNL2* as a possible risk gene in this study is positive and validates the approach taken in using linkage analysis of extended pedigrees as it had already been reported as associated with IOP. *HERC2* is a novel gene for IOP identified in our current study, and very recently the locus has been (independently) identified by GWAS. *PKP4* is also a novel gene for IOP and POAG. *PKP4* and *HERC2* require further research in relevant ocular tissues to determine whether they are functionally important and whether they could affect IOP regulation.

This is one of the first studies to conduct linkage analysis using whole exome sequencing data on intermediate traits of POAG. This research has identified three candidate genes which may be involved with IOP regulation. Currently as the only modifiable risk factor for POAG, IOP is an ideal target to develop more effective treatments than are currently available. This research is necessary to improve our understanding of IOP regulation and its involvement with POAG pathogenesis.

Abbreviations

alt	alternate allele
bp	base pairs
BED file	browser extensible data - file format
CCT	central corneal thickness
CDR	cup to disc ratio
Chr	Chromosome
cM	centiMorgan
CNV	copy number variant
DNA	deoxyribonucleic acid
ESE site	exonic splicing enhancer site
ESS site	exonic splicing silencer site
GATK	Genome Analysis Toolkit
GIST	Glaucoma Inheritance Study in Tasmania
GWAS	genome-wide association study/studies
HTG	high tension glaucoma
HCNC	HUGO gene nomenclature committee
IBD	identity by descent
IBDLD	program used to estimate IBD
IOP	intraocular pressure
JOAG	juvenile open-angle glaucoma
LD	linkage disequilibrium
LOD	logarithm of the odds
MAC	minor allele copies
MAF	minor allele frequency
MVIP	Melbourne Visual Impairment Project
NGS	next generation sequencing
NTG	normal tension glaucoma
ODA	optic disc area
OGGS	Oregon Glaucoma Genetics Study
OMIM	online Mendelian inheritance in man
OTD	ocular tissue database
PACG	primary angle-closure glaucoma
PCG	primary congenital glaucoma
POAG	primary open angle glaucoma
PRS	polygenic risk score
QTL	quantitative trait locus
ref	reference allele
RGC	retinal ganglion cells
SOLAR	Sequential Oligogenic Linkage Analysis Routines program
SNP	single nucleotide polymorphism
UTR	untranslated region
VCDR	vertical cup to disc ratio
VCF	variant call format – file format
WES	whole exome sequencing
WGS	whole genome sequencing

Contents

Tables

Figures

Chapter 1
Introduction

1.1 Introduction to glaucoma

Glaucoma encompasses a group of eye diseases which result in damage to the optic nerve and vision loss, and is the leading cause of irreversible blindness worldwide [1, 2]. Glaucoma is classified into subtypes based on the age of onset of the disease, the structure of the anterior chamber of the eye and whether there are other underlying causes. Secondary glaucoma is caused by some medications, trauma to the eye or other eye diseases, such as pseudoexfoliation syndrome [3]. Primary forms of glaucoma have no other underlying causes. Primary congenital glaucoma (PCG) is a serious condition of infancy, resulting from a developmental defect in the eye [4]. Juvenile open-angle glaucoma (JOAG) and primary open-angle glaucoma (POAG) are distinguished by their age of onset, with JOAG developing between the ages of 3 and 40 years and POAG after the age of 40 years [2, 5]. The "open-angle" refers to a normal angle size between the iris and cornea (Figure 1-1a). Primary angle-closure glaucoma (PACG) is characterised by an abnormal position of the iris, closing the angle between it and the cornea (Figure 1-1b) [3]. Patients with PACG have a physical obstruction to the outflow of aqueous humour from the anterior chamber of the eye. Those with POAG do not have the same physical obstruction, although they may have resistance to aqueous humour outflow caused by other means [3]. POAG is the most common form of glaucoma, making up approximately 75% of glaucoma cases [6].

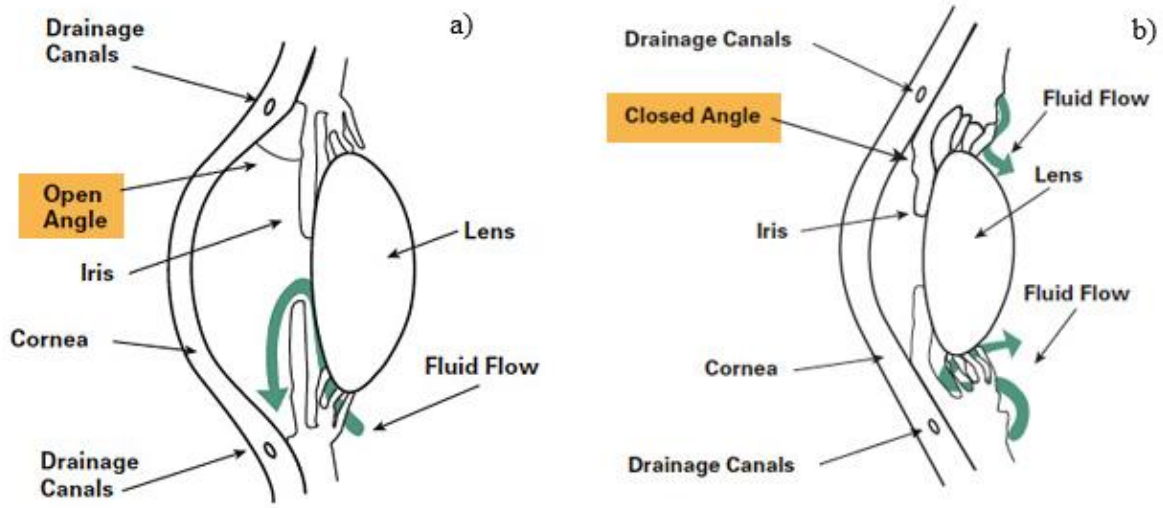


Figure 1-1 Cross sectional diagrams of the anterior chamber of the eye showing a) a normal open-angle and b) a closed angle. Aqueous humour fluid flow is shown as a green arrow. Source: http://www.glaucoma.org/glaucoma/types-of-glaucoma.php

1.2 Prevalence

PCG and JOAG are rare forms of glaucoma and their prevalence is difficult to estimate. POAG and PACG account for the majority of glaucoma cases worldwide. In 2010, it was estimated that 8.4 million people worldwide were bilaterally blind due to POAG or PACG [2, 6]. With an aging worldwide population, this estimate has risen to 11.1 million in 2020 [6]. POAG, on its own, currently affects approximately 55 million people [1, 6, 7] and this is projected to rise to nearly 80 million by 2040 [1]. Ethnicity affects the prevalence of glaucoma quite dramatically. PACG is more prevalent in Asian populations (at approximately 1%) and least prevalent in North Americans (< 0.3%) [1]. While approximately 2% of over 40 year olds worldwide have POAG, this figure rises to over 4% in those of African heritage [1, 7]. In Australia, the prevalence of those with POAG is similar to the worldwide average. In 2020, between 200,000 and 300,000 Australians are estimated to be affected by this disease [7, 8]. Interestingly, Indigenous Australians have a much lower prevalence of glaucoma than the non-Indigenous population (1.6% versus 3.4% respectively), even though the Indigenous population have more ocular risk factors, such as large optic heads and thinner central corneal thickness [8]. With such large numbers of people affected by POAG, the economic burden on the health system is great [9]. More importantly, the quality of life decreases for those with a visual impairment [9]. As the incidence of POAG increases with age, the elderly are

particularly at risk of diminished quality of life. With visual impairment, they risk tripping and falling over, losing their independence and increased rates of depression [9]. With an ageing worldwide population, the effects of POAG on both economic and personal levels will only increase.

1.3 Pathophysiology of POAG.

POAG is itself a heterogeneous disease, with a range in the age of onset, amount of optic nerve damage, amount of visual field loss, whether the pressure inside the eye is raised and response to treatment. Although individuals are affected by POAG differently, there are characteristic features of this disease.

1.4 The optic nerve and optic disc

The innermost layer of the retina is comprised of retinal ganglion cells (RGC, Figure 1-2). The axons of the RGC converge to form the optic nerve, which transmits visual signals to the lateral geniculate nucleus, from where information is relayed to the visual cortex of the brain [10]. In a healthy eye, there are approximately one million axons in an optic nerve [11]. The optic disc is the raised disc on the retina at the point where the optic nerve leaves the eye. The optic cup is the depression in the middle of the disc, where the axons converge. This lighter coloured cup is surrounded by a pink neuroretinal rim [12]. Together, the cup and rim make up the optic disc. Characteristic changes in the optic disc are used for diagnosis and evaluation of the glaucomatous process.

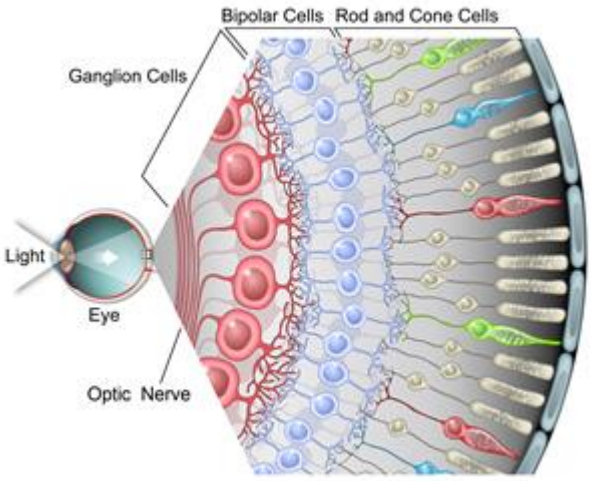


Figure 1-2 Cell types of the retina The retinal ganglion cells (RGC) are in the innermost layer. The axons of the RGC converge to form the optic nerve. Source: http://www.salk.edu/news-release/from-eye-to-brain-salk-researchers-map-functional-connections-between-retinal-neurons-at-single-cell-resolution/

1.4.1 Optic nerve damage in POAG

POAG is a chronic and progressive neurodegenerative disease, affecting the optic nerve of the eye [6]. POAG generally affects both eyes, but is often asymmetrical, and one eye may be more severely affected than the other [11]. In POAG, there is progressive apoptosis of the RGC with a resultant loss of vision. This cell death occurs in a characteristic pattern resulting in loss of axons which pass through the upper and lower poles of the optic disc [2]. Mid-peripheral vision is the first to be affected. Central vision is not affected until much later in the glaucomatous disease process [2, 13]. Death of RGC do not just affect the eye. Disruption of axonal transport in these cells also affects the visual centres of the brain [10]. RGC also rely on neurotrophic support provided by their brain target neurons as well as their retinal interactions. As RGC apoptosis begins, the neurotrophic support diminishes, triggering further RGC death [2].

1.4.2 Vertical cup to disc ratio changes in POAG

Along with vision loss, changes in appearance of the optic disc provide a clear indication of glaucomatous damage. As RGCs die, there is a characteristic “cupping” of the centre of the disc. Loss of RGC and deformation of connective tissues cause the cup to become deeper and wider. The neuroretinal rim surrounding the cup becomes thinner and paler. There may be other structural changes such as disc haemorrhages and notching of the neuroretinal rim [2, 10, 11]. An important clinical measurement which is used, to determine level of damage, is the cup to disc ratio (CDR). This measurement is generally taken in a vertical orientation (VCDR) (Figure 1-3a). The diameter of the cup divided by the diameter of the whole disc gives a ratio which indicates the amount of glaucomatous damage. The more damage there is due to cell death, the larger the cup, which leads to a larger ratio (Figure 1-3b). A healthy eye has a VCDR of approximately 0.3 and glaucomatous eye is generally considered to be ≥ 0.7 [14]. Asymmetry of VCDR of > 0.2 between the two eyes may also be an indication of RGC damage [11]. As there is variation in disc size and VCDR in a normal, healthy population, VCDR alone is not relied upon to determine if there is glaucomatous damage [14].

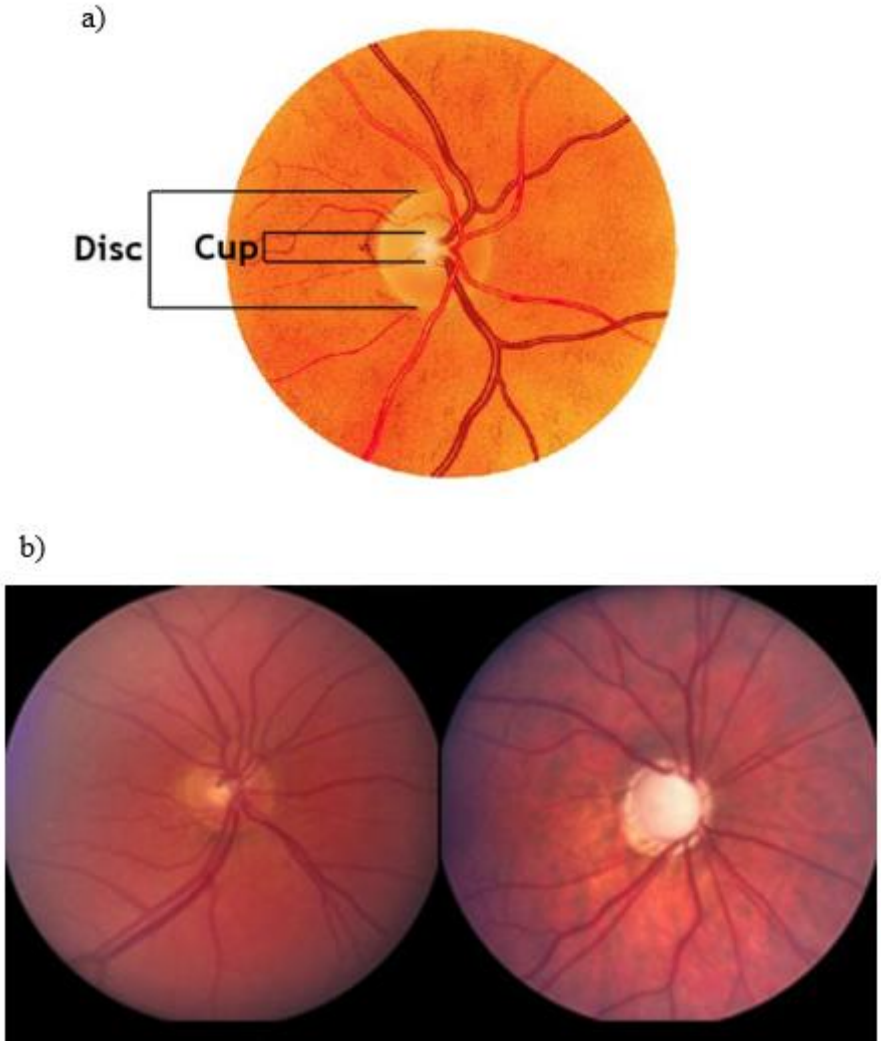


Figure 1-3 Photographs of the back of the eye showing optic discs a) Vertical cup to disc ratio (VCDR) is measured by dividing the diameter of the cup by the diameter of the disc in the vertical orientation. b) A comparison of the difference in appearance of a normal optic disc (left) with a glaucomatous optic disc (right), which has an enlarged cup and thinner neuroretinal rim.
Sources: a) http://www.intechopen.com/books/the-mystery-of-glaucoma/the-optic-nerve-in-glaucoma b) http://www.glaucoma.org/treatment/optic-nerve-cupping.php

1.4.3 Intraocular pressure

Many patients with POAG have an elevated intraocular pressure (IOP), which is the pressure inside the eye. Generally, a pressure greater than 21mmHg, as measured by tonometry, is considered to be high [11]. In these patients, there is often a resistance to aqueous humour outflow from the anterior chamber of the eye. In a normal eye, there is a balance of aqueous humour secretion by the ciliary body and its drainage through the trabecular meshwork and uveoscleral outflow pathways [3] (Figure 1-4), resulting in a normal pressure of between 11mmHg and 21mmHg [10]. However, in some patients, the trabecular meshwork does not

drain adequately. The mechanism for this resistance is not fully understood, but may involve a loss of lining cells or a blockage within the trabecular meshwork [2]. This results in an increase in pressure in the anterior chamber of the eye, which causes an increase in pressure in the vitreous humour in the posterior chamber, where the optic nerve is located [2]. It is possible that increased pressure is responsible for optic nerve damage due to mechanical stress [10, 15]. The mechanical stress may affect RGC, the optic disc itself, or other tissues such as the lamina cribosa, the tissue which the optic nerve passes through to exit the eye [3]. It has also been suggested that an increase in IOP causes an accumulation of reactive oxygen species in the eye, and this causes oxidative damage to the RGC [10, 15]. It has still not been determined as to which aspect of high IOP is important for progression to POAG; mean IOP, maximum IOP or fluctuations in IOP, which occur normally over the course of the day [13, 16].

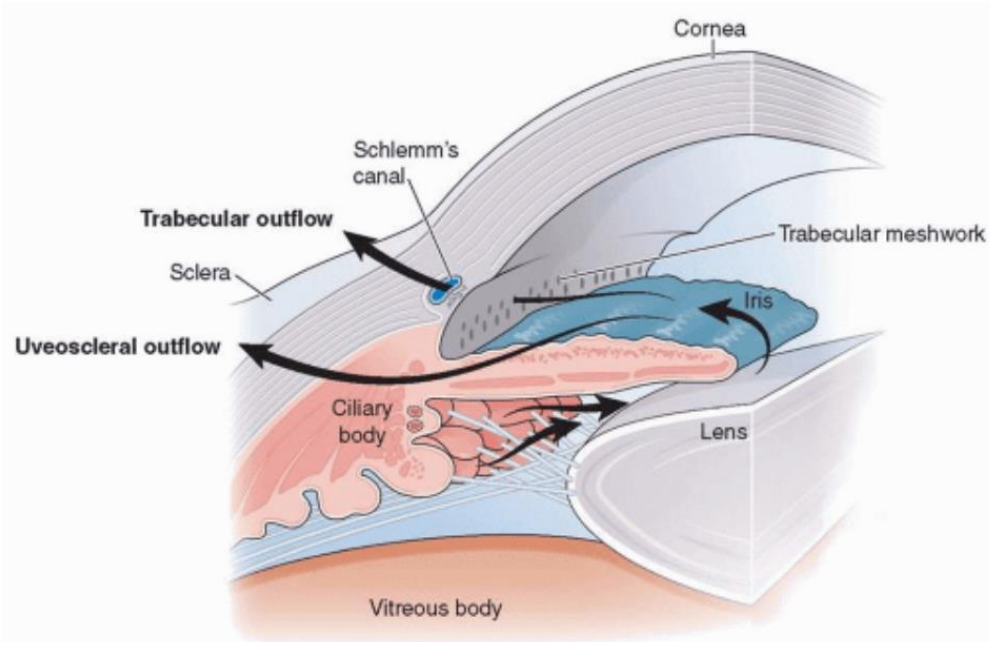


Figure 1-4 Aqueous humour outflow through the eye. In a healthy eye, intraocular pressure is maintained by aqueous humour production in the ciliary body and drainage via the trabecular meshwork and uveoscleral pathways.
Source: https://entokey.com/anatomy-of-the-aqueous-outflow-pathways/

Although most people with POAG do have an elevated IOP, a large percentage (30 – 40%) do not [2, 10, 17]. These patients have a normal IOP (≤ 21mm Hg), but similar optic nerve damage and visual field losses as those with higher IOP. This is referred to as normal tension glaucoma (NTG), to reflect that fact that there is glaucomatous damage with "normal" eye pressure. Sometimes, NTG is discussed as though it is a separate form of glaucoma to POAG [13, 18].

Others make the distinction between NTG and high tension glaucoma (HTG) as categories within POAG [19]. Often, NTG is included in the general POAG grouping, and projections of POAG prevalence worldwide do not make the distinction between NTG and HTG [1, 6, 7]. Regardless of how NTG is categorised, it is more common in some ethnic groups than others, for example, Japanese have higher prevalence than Caucasians [19, 20]. Aside from patients with NTG having optic nerve damage with lower IOP, they also have greater visual field loss, for the amount of optic nerve damage, than those with HTG [10]. It has been suggested that optic nerve damage may not so much be the result of having a high IOP, but rather the difference in pressure between the eye and the brain. This pressure difference, the translaminar pressure difference (TPD), is a result of the lamina cribosa tissue separating the pressurised eye from the cerebrospinal fluid pressure in the brain [10, 13, 21]. The role of TPD is still not clear, as it is difficult gaining accurate intracranial pressure measurements via non-invasive means [21]. There are indications that patients with NTG may have lower intracranial pressure resulting in a higher TPD than healthy individuals. Also, NTG patients may be more susceptible to nerve damage at lower pressures than those with HTG [21].

1.5 Diagnosis of POAG

An early diagnosis of POAG is key to slowing the progression of the disease. Currently, 50% of those affected by POAG in developed countries, including Australia, are undiagnosed. This figure rises to 90% on a worldwide level [2, 6]. As POAG is initially asymptomatic, painless and progresses slowly, diagnosis may not occur until there is substantial vison loss [22]. The patients themselves may not notice a loss in vision until the disease is quite advanced, by which time the optic disc may be severely damaged [13]. It is estimated that between 30-50% of RGC could die before any obvious vision loss has taken place [11, 23], hence the importance of an early diagnosis of optic nerve damage.

To be diagnosed with POAG, a patient needs to be examined by a clinician for optic nerve damage and visual field loss. As there is a wide variation in optic disc size and VCDR in the population, the importance of repeated measurements taken over time becomes apparent. An increase in VCDR may be an indication of early stage optic nerve damage, which may possibly be halted before vision loss has occurred. Visual field testing is conducted using perimetry, which is often automated, and the clinician looks for specific visual field defects characteristic of glaucoma [23]. As perimetry is a subjective test, dependent on patient responses, repeated

measurements over time are also important to ensure consistency and to determine the progression of vision loss. It is important that VCDR measurements and perimetry results are concordant to diagnose POAG, as vision loss alone may be due to other causes [23].

Although IOP is not an essential diagnostic criterion for POAG, most patients with POAG do have increased eye pressure [2]. Tonometry is routinely used at optometrists, as part of a regular eye examination, to measure elevated pressure levels in the eye. An IOP of ≥ 22mHg may not necessarily mean that optic nerve damage has occurred, however, referral to an ophthalmologist may be indicated to reduce the risk of developing POAG.

1.6 Treatment of POAG

Once POAG has been diagnosed, the only proven method to treat it is to reduce the pressure inside the eye [24]. Even in NTG patients, who do not have an elevated IOP, reduction in pressure has been shown to reduce visual field loss [10, 25]. For patients with high IOP, but no glaucomatous damage, administering eye drops can reduce the risk of POAG-related vision loss [26].

There are several different types of eye drops which are usually the first treatment provided to reduce IOP. Some work by increasing outflow of the aqueous humour from the eye and others work by decreasing production of aqueous humour [3]. Either way, the purpose is to reduce the amount of aqueous humour in the anterior chamber of the eye, to reduce the pressure on the optic nerve in the rear of the eye. Although eye drops are well tolerated and with only few side effects, patient adherence with regular and continued administration may be an issue [2]. Even with regular administration, eye drops alone may not be adequate to halt the progression of the glaucoma. There are surgical options to increase aqueous humour outflow and reduce IOP; trabeculoplasty is a laser surgery which enlarges drainage channels out of the eye and trabeculectomy is a traditional surgical technique which removes part of the trabecular meshwork of the eye [10]. The currently available glaucoma treatments are not able to cure the disease, but are able to slow down the progression of optic nerve damage and preserve the remaining vision to maintain the quality of life for the patient [3]. Neuroprotective treatments would be beneficial to protect the optic nerve head from damage, however, as the pathophysiology of POAG is not well understood, it is currently difficult to develop such therapies [3].

1.7 Risk factors for POAG

There are several risk factors which increase an individual's chance of developing POAG. Boland and Quigley (2007) [16], provide a comprehensive list of possible risk factors based on different categories, such as state of the individual (age, sex and ethnicity), systemic diseases and personal behaviours. That study, as well as many others, agree that increased IOP is the main and only modifiable risk factor [10, 11, 15, 16, 27-29]. Although an increased IOP is not always a symptom of POAG, it is one of the most important risk factors and is the only factor which can be treated.

Age is another important risk factor for the development of POAG. POAG is an adult onset disease, usually developing after the age of 40 years [7]. The Blue Mountains Eye Study suggests an exponential increase in the prevalence of POAG after the age of 60 years, with less than 0.5% developing the disease before the age of 60 years to over 8% for those aged over 80 years in this population [30]. The increasing of risk as people age may represent a measure of various factors, such as length of time exposed to other risk factors, or the deterioration of ocular tissues over time, making them more susceptible to glaucomatous damage [16].

Ethnicity is another important risk factor for developing POAG. People of African heritage develop POAG earlier, do not respond as well to treatments and have a much higher prevalence than Caucasians [6, 7, 10, 16]. Over 6% of those with an African heritage will have developed POAG by age 65, as compared to only 2% of Caucasians [7].

There are also ocular risk factors for the development of POAG. High myopia (short sightedness) is one of these [11, 16]. Central corneal thickness (CCT) has also been proposed as a risk factor, with the thinner the CCT, the greater the risk of developing POAG [11, 13]. However, CCT also affects IOP measurements. IOP is most commonly measured with the Goldmann applanation tonometer [31], which may provide inflated IOP measurements in individuals with thicker corneas and lower IOP levels in those with thinner corneas. Consequently, patients with thicker corneas and higher measured IOP levels are prescribed pressure reducing medication more than those with thinner corneas and lower IOP measurements. As a result, patients with thinner corneas, who are less medicated, are associated with increased risk of developing glaucoma [13, 16]. CCT is a highly heritable ocular measure [32], and although it may affect IOP measurements, is still believed to be a risk factor for POAG in itself [33].

A large optic disc area (ODA) has also been proposed as a risk factor [2, 16]. The larger discs may be more susceptible to nerve damage, although they may potentially contain more nerve fibres [16]. Optic cup area and CDR have also been proposed as risk factors for POAG [3, 11], however, they are also measures of optic nerve degeneration. Regular examination of patients with large optic cups and large CDR, without POAG, will help to determine the amount of risk these factors confer [16].

Other POAG risk factors have been assessed, such as; smoking, diabetes, body mass index, sex, hypertension, alcohol consumption and dietary fat intake. Many of these studies have resulted in conflicting and inconclusive results, as reviewed in [13, 15, 18, 27, 34-37], and even if some of these prove to be true POAG risk factors, they are unlikely to be as important as the other risk factors previously discussed.

1.8 Family history of POAG

A family history of POAG is a well-recognised and a very important risk factor for the disease. Although it often signifies shared environmental background, the family history also represents the genetic background [16]. It has been demonstrated there is a ten-fold increase in the risk of POAG if an individual has a close family member with the disease [38]. A study by Wolfs et al. (1998) [38], found that a parent or sibling with POAG increases the chance of developing the disease to 22% as compared with 2.3% in relatives of non-glaucoma controls. The risks may actually be underestimated if the children were too young to have developed POAG at the time of the study [22]. Gong et al. (2007) [39], distinguish between inherited and familial forms of POAG and suggest that these two forms account for 72% of POAG cases. They define inherited POAG as exhibiting a Mendelian inheritance pattern with high penetrance and familial POAG exhibiting more complex inheritance patterns with reduced penetrance. Sporadic POAG cases are those single patients without affected first or second degree relatives. As familial data continues to be collected over time and more information is gathered, the definition of the families may change. Classifying POAG into these categories may guide genetic studies as well as allowing for those at greater risk of developing the disease to undergo appropriate genetic testing [39].

The heritability of POAG in three different European ancestry populations, ranged from 42% to 81% [40-42]. These heritabilities estimate the amount of phenotypic variation in POAG which is due to the genotypic variation. The higher heritability estimate was from a study on

large, extended pedigrees [42] and the lower estimate was from a genome-wide association study (GWAS) [40] with a twin study providing a heritability estimate mid-way between the two [41].

Several studies have also determined the heritability of some of the intermediate traits involved with POAG, such as; intraocular pressure, central corneal thickness, optic disc diameter and vertical cup to disc ratio [27, 42-45]. The heritability for these traits are also high, demonstrating the importance of genetics in POAG. There is little doubt that genetic background is an important risk factor in this disease, at the family level through to the population level. Different populations, as represented by different ethnicities, display varying prevalence to POAG [7] .Within each ethnic group, there is also the genetic background of each family, which is involved in determining an individual's risk of developing POAG.

1.9 Genetics of POAG

Over the last 25 years or so, there has been a wealth of research centred around determining the genetic causes of POAG. The use of two distinct approaches has resulted in many genetic loci being proposed in POAG pathogenesis. POAG is a genetically complex disease, with both causative genes and susceptibility genes identified.

1.9.1 POAG genes and loci identified by family-based studies.

Family-based studies identified the first genetic loci linked to POAG. Since the first glaucoma locus, GLC1A, was identified in 1993 [46], over 20 more glaucoma loci have been found through the use of linkage analysis in families [13, 27]. The major loci are named GLC1A – Q and there are several other linkage regions which are un-named [28]. The HUGO gene nomenclature committee (HGNC, https://www.genenames.org/) has withdrawn the "phenotype only" locus nomenclature and no longer recognises these loci without identified causative genes. The online Mendelian inheritance in man catalogue (OMIM, https://omim.org/) does still include these loci. In this dissertation these loci will be referred to by their original GLC1A - Q nomenclature as a convenient way to refer to previously identified genomic locations linked to glaucoma which may not yet have identified causative genes. Loci GLC1J, K, M and N are linked with juvenile open-angle glaucoma [47-49]. A locus originally

identified in primary congenital glaucoma, and named GLC3A [50], is also recognised as being important in the adult onset form of the disease [13, 28].

These linkage analyses resulted in the identification of large linkage regions, spanning millions of base pairs, and it was difficult to identify causative genes. Deeper sequencing was necessary to identify if candidate genes within these regions carried deleterious, potentially disease causing variants. Seven genes have been identified from within these linkage regions and these account for between 5 and 10% of all POAG cases [18, 51, 52]. Considering that the heritability of POAG is so high [39, 42], the influence of these genes at a population level, may not be so important. However, at the family level, the identification of POAG causative genes has allowed family members to assess the risks of developing the disease, undergo genetic testing for at risk members and have an earlier intervention with treatments [53]. One of the most important genes identified from linkage analysis is myocilin (*MYOC*), and testing for *MYOC* mutations has demonstrated that at risk individuals can be identified before signs of glaucoma are apparent [54]. The identification of POAG causative genes and the biological pathways they influence also enable a greater understanding of glaucoma pathogenesis, regardless of the process of gene identification [13, 37]. The seven genes which have been identified from linkage regions, using family-based studies, are summarised in Table 1-1. Four of the most important of these; *MYOC*, *OPTN*, *TBK1* and *CYP1B1* are discussed in more detail below.

GLC1A was originally identified in a large family with JOAG [46]. The *TIGR* gene was identified from within this locus in 1997 [55], subsequently renamed *MYOC*, and has since become one of the most widely studied glaucoma genes. There is a dedicated online database for *MYOC* (http://www.myocilin.com/) [56], which allows researchers to interrogate specific mutations within the gene. Over 180 variants have been described in *MYOC.* Around 85% of these are missense mutations and 40% of variants have been identified as disease causing [56]. The majority the variants are in the third exon of the gene [15, 57]. The Gln368STOP mutation is the most common *MYOC* variant and is associated with a later onset form of POAG, with a diagnosis after the age of 35 years [58]. *MYOC* variants are more prevalent and cause a more severe form of glaucoma in JOAG cases (22 – 36%) than in POAG (2 - 4%) [19, 28, 53].

Increased IOP is a feature of *MYOC* mutations. *MYOC* is expressed in all ocular tissues, including the trabecular meshwork. It has been proposed that variants in the gene affect protein folding and prevent the protein from being secreted, causing accumulation in the trabecular meshwork, resulting in increased IOP [15, 19, 28, 53, 57]. Although reduction in pressure is

currently the only form of treatment for patients with *MYOC* mutations, research using transgenic mice holds promise for targeted treatments which may prevent the elevation of IOP in the first place [53].

The second gene to be identified from POAG linkage regions was *OPTN* in the GLC1E locus [59]. Variants in *OPTN* are associated with NTG. These mutations account for less than 2% of NTG cases [19, 28, 57]. *OPTN* is expressed in several ocular tissues, including the trabecular meshwork and retina. The most common mutation, E50K, causes a more severe form of glaucoma and at a younger age of onset. Mutations in the *OPTN* gene may cause abnormal protein deposits in the retina and apoptosis of RGC [60].

TBK1, in the GLC1P locus, was identified from family studies of NTG [61]. Unlike *OPTN*, in which coding variants cause disease, the *TBK1* mutation is a copy number variation. Duplications and triplications of *TBK1* are responsible for around 1% of NTG cases [57, 62]. *TBK1* is expressed in the retina and has been shown to bind to *OPTN* and act in the same biological pathway [63, 64].

The GLC3A locus was identified in families and linked to PCG [50]. Since then, and with the identification of the *CYP1B1* gene within this locus, it is now recognised as an important gene in the development of JOAG and POAG as well as PCG [65-67]. *CYP1B1* is expressed in several tissues of the eye and mutations in this gene may be responsible for developmental abnormalities [66]. *CYP1B1* interacts with *MYOC* and some carriers with mutations in both these genes exhibited an earlier age of onset and more severe form of glaucoma [67].

.

Table 1-1 Summary of important findings for the seven genes identified from linkage studies of POAG families.

Gene (Locus) Name	References	Inheritance pattern	Function of gene / involvement in POAG	Reviewed in:
MYOC (GLC1A) myocilin	**Locus**: Sheffield el al., 1993 [46] **Gene**: Stone el al., 1997 [55]	AD	• Encodes a secreted glycoprotein which is expressed in all cells including the trabecular meshwork and other ocular tissues. • Abnormal myocilin protein may be toxic to trabecular meshwork cells. • Many mutations found in gene (>180), ~ 40 are disease causing. GLN368STOP is the most common mutation. • High IOP a feature of a *MYOC* mutation. • Accounts for 2-4% of POAG and 22-36% of juvenile glaucoma.	[13, 15, 19, 28, 37, 51, 57, 68, 69]
OPTN (GLC1E) optineurin	**Locus**: Sarfarazi el al., 1998 [70] **Gene**: Rezaie el al., 2002 [59]	AD	• Function of protein implicated in tumour necrosis factor-alpha pathway. • Expressed in many ocular tissues. • Associated with NTG. • May be involved with RGC death.	[13, 15, 19, 28, 37, 51, 57, 68, 69]
ASB10 (GLC1F) ankyrin repeat- and socs box-containing protein 10	**Locus**:Wirtz el al., 1999 [71] **Gene**:Pasutto el al., 2012 [72]	AD	• Expressed in several tissues of the eye including RGC, ciliary body and trabecular meshwork. • May be involved in IOP regulation.	[15, 19]
WDR36 (GLC1G) WD repeat-containing protein 36	Monemi el al., 2005 [73]	AD, complex	• Encodes a protein with unknown function which is expressed in all cells. • Controversy over the role of WDR36 in glaucoma due to inconsistency of results in different studies. • May need the involvement of mutation in *STI1* to have functional consequence in glaucoma.	[13, 15, 19, 28, 37, 51, 57, 69]
NTF4 (GLC1O) neurotrophin 4	Pasutto el al., 2009 [74]	complex	• Neurotrophic factor important for neuron survival. • Variants in this gene are rare. • Inconsistent results in different studies. • May be involved with RGC death with high IOP.	[15, 28, 57, 68]
TBK1 (GLC1P) TANK-binding kinase 1	**Locus:** Fingert el al., 2011 [61] **Gene:** Fingert et al., 2014 [75]	complex	• Copy number variation associated with POAG. Duplication of gene associated with NTG. Gene triplication also found associated with POAG. • Expressed in ocular tissues. • Interacts with *OPTN.*	[28, 37, 51, 57, 62]
CYP1B1 (GLC3A) cytochrome P450, subfamily1, polypeptide 1	Sarfarazi el al., 1995 [50]	AR	• Expressed in several tissues of the eye. • A major locus involved with juvenile glaucoma. • Possibly interacts with *MYOC* in POAG.	[15, 19, 28, 51]

AD = autosomal dominant, AR = autosomal recessive, POAG = primary open angle glaucoma, NTG = normal tension glaucoma, IOP = intraocular pressure, RGC = retinal ganglion cells

1.9.2 POAG genes and loci identified by population-based studies

While family-based linkage studies have identified POAG causative genes, population-based studies have proposed genes which may be involved with increased susceptibility to POAG. GWAS have proven an effective tool in identifying single nucleotide polymorphisms (SNPs) which are associated with POAG. The first POAG GWAS was published in 2009 [76], and identified six SNPs associated with POAG, although these SNPs did not reach genome-wide significance at $p < 5x10^{-8}$. The *CAV1*/*CAV2* locus was the first to be associated with POAG at genome-wide significance and replicated in independent cohorts [77]. Population-based candidate gene studies have also proven popular and have replicated findings from GWAS and found associations with novel POAG genes [27]. Population-based studies at both the genome-wide and candidate gene level have been responsible for the identification of more than 74 loci which have reached genome-wide significant association with the POAG phenotype [13, 27, 28, 78].

Although many loci have been proposed through GWAS, the genes which may be responsible for POAG pathogenesis are still difficult to identify and validate. The majority of associated SNPs from GWAS are in non-coding and intergenic regions and usually the nearest gene to that SNP is proposed as the associated gene [79]. These proposed genes need to be validated in other studies to confirm their involvement in POAG pathogenesis. The alleles of genes identified in GWAS may increase the risk of developing POAG, but do not cause the disease on their own. Many of these risk alleles and/or particular combinations of them are required to give an individual a higher susceptibility of developing POAG. There are also likely to be gene-gene and gene-environment interactions of these risk alleles, adding to the complex genetic architecture of POAG [51]. While predictive genetic testing is now available for the some of the POAG causative genes [53, 80] the complexity of the risk alleles means that predictive testing for susceptibility to POAG is still in its early stages. In more recent times, the use of polygenic risk scores (PRS) to predict an individual's risk of developing a disease, is becoming a popular method of using GWAS generated data [81]. The use of PRS in complex diseases has great potential in the clinical setting, allowing those individuals at greater risk for a particular disease, increased monitoring and earlier intervention if necessary [82]. PRS analysis for glaucoma is a very recent and exciting development for understanding which SNPs contribute to an increased risk of developing this disease. PRS analysis using glaucoma risk loci [83] and IOP loci [84, 85]

have only recently been conducted and provide promising research in identifying at risk individuals.

A sample of POAG susceptibility genes, which have been replicated in several population based studies, are listed in Table 1-2. Although these loci are all associated with the POAG phenotype, it is not yet understood how these genes are involved in the biological pathways which lead to POAG pathogenesis.

Table 1-2 POAG susceptibility genes from population based studies. These are highly replicated studies and represent a sample of possible POAG susceptibility genes.

Nearest gene	SNP	References
CAV1/CAV2 caveolin 1/caveolin 2	rs10258482	[86]
	rs10262524	[86]
	rs1052990	[77, 87]
	rs17588172	[88]
	rs4236601	[77, 87-89]
	rs7795356	[90]
CDKN2B-AS1 CDKN2B antisense RNA 1	rs10120688	[91-94]
	rs1063192	[29, 69, 90, 95-99]
	rs1412829	[91, 95]
	rs2157719	[94, 100]
	rs4977756	[89, 91, 94, 95]
	rs523096	[101, 102]
	rs7049105	[92, 94]
	rs7865618	[101, 103]
SIX1/SIX6 sine oculis homeobox, drosophila, homolog of, 1/6	rs10483727	[29, 69, 89, 90, 94, 95, 97-99, 104-106]
	rs4901977	[103]
	rs146737847	[107]
	rs33912345	[107, 108]
TMCO1 transmembrane and coiled-coil domains 1	rs4656461	[86, 89, 91, 95, 109]
	rs7518099	[86, 91, 110]
	rs7555523	[66, 86]

Nearest gene = the closest gene to the SNP as identified in the original paper.

Janssen et al. (2013) [13], incorporated 65 POAG associated genes in a biological pathways analysis (including the genes in Table 1-2 and the causative genes in Table 1-1), and predicted four different molecular networks. The genes from Table 1-2 were divided across all four networks, including the separation of the *SIX1* and *SIX6* genes into two different pathways. *SIX6* was linked with *OPTN* in a network which included the function of "visual system development".

The pathway which included the function of "ophthalmic disease", incorporated the *MYOC* and *CDKN2B-AS1* genes. The other two pathways did not explicitly state a visual function. The four pathways demonstrated intricate webs of interactions of the POAG associated genes, underlying the complex genetic nature of this disease.

1.9.3 Missing heritability in POAG

Although GWAS have been very successful in proposing POAG associated loci, in total these loci only account for around 3% of the heritability of this disease [111]. These are common variants each with small effect sizes. The genes identified from linkage studies are rare variants with large effect sizes and account for 5-10% of POAG cases [27, 112]. Clearly there is a great deal of "missing heritability" for this complex disease. There have been many suggestions proposing the reasons as to why the heritabilities of identified variants do not add up to the total trait heritability, which for POAG is up to 80% [42]. GWAS may identify many associated loci, but they have small effect sizes and do not account for a large amount of the total heritability [113-115]. Genotyping arrays used for GWAS use common SNPs and are designed for large case-control cohort studies. Structural and copy number variants also have an impact on disease, but GWAS are underpowered to find them with statistical significance [113, 114, 116]. The *TBK1* CNV found in NTG patients (discussed in section 1.9.1 above), was identified from linkage analysis [61, 75]. Gene-gene interactions and gene environment interactions are also likely to be important in disease risk and progression, as are epigenetic effects [114, 116, 117]. These are all likely to account for some of the missing heritability, but none of these are able to be assessed with GWAS . Importantly, rare variants with larger effect sizes may account for a substantial amount of this missing heritability [113, 117-119]. Figure 1-5 shows the differing effect sizes of rare and common variants and that different experimental approaches need to be taken to identify them. GWAS are useful in identifying many variants with small effect sizes, but different approaches need to be taken to identify rare variants with larger effects. Wainschtein et al. (2019) [120], have demonstrated recovery of the gap in heritability when rare variants were included from whole-genome sequencing (WGS) data, which were missing from the imputed data from SNP arrays. In that study, more than half of the heritability for height was accounted for by variants with a minor allele frequency (MAF) between 0.0001 to 0.1. Many of these variants were rare and in low linkage disequilibrium (LD) with other genomic variants and would be impossible to find using traditional genotyping arrays [120].

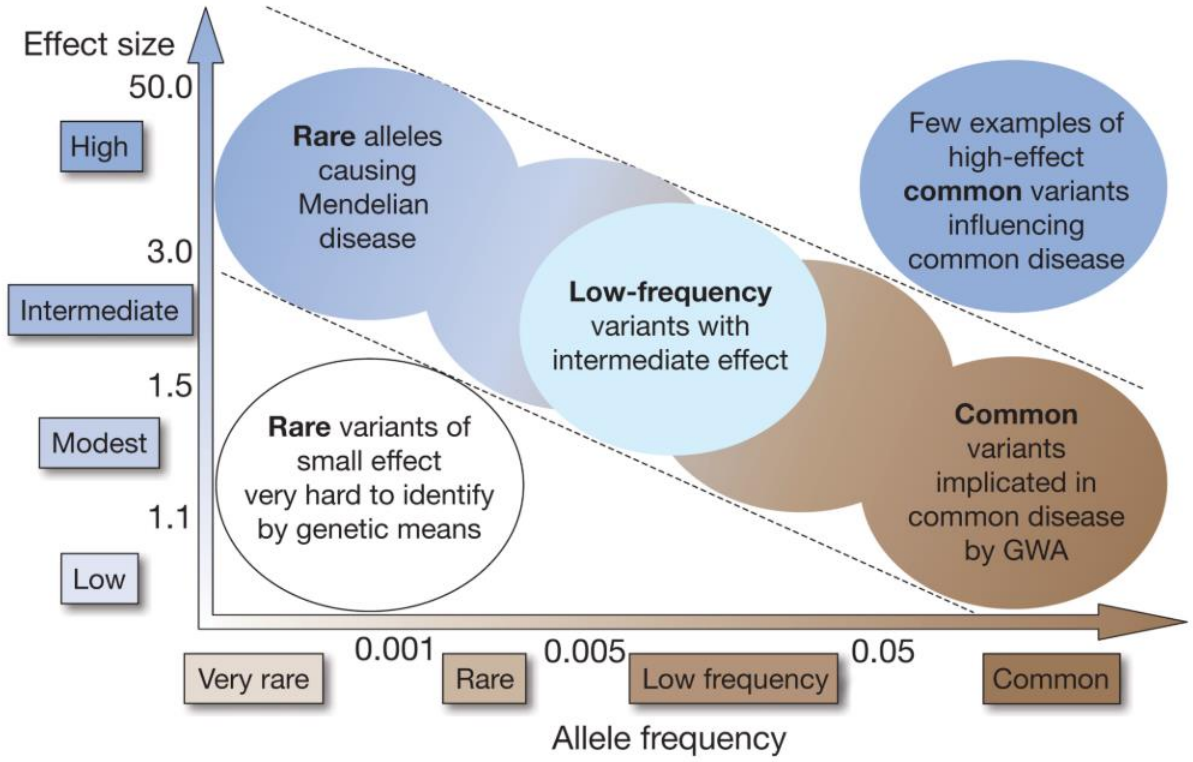


Figure 1-5 Identifying genetic variants based on risk allele frequency and genetic effect Diagonal lines represent where the most emphasis is on identifying associations. Source: Manolio et al. (2009) [113]

The use of linkage analysis is an appropriate method to identify less common variants with large effect sizes. It has traditionally been used to identify genes involved with Mendelian disorders, diseases in which a single, highly penetrant mutation is causative. Although successful in identifying the few, rare POAG causative genes (as discussed in section 1.9.1 above) linkage is not commonly used to find genes involved in complex diseases due to the multiple underlying genetic effects [121]. POAG is a highly complex, late onset disease and it is difficult obtaining families with multiple generations, as parents and siblings may have already died by the time family members are diagnosed [18, 28, 52]. As the importance of rare variants in complex diseases is becoming better understood [122-124], linkage analysis provides a better tool for identification of these variants than GWAS, which is suited to common variants. Families provide improved statistical power to search for rare variants affecting disease due to the natural enrichment of these variants as they segregate through the generations [125, 126]. Large families with many family members affected by POAG provide a valuable resource for finding rare variants involved with this disease.

1.10 Quantitative clinical traits of POAG

Part of the missing heritability for POAG, discussed above, may be accounted for by identifying variants influencing the quantitative, intermediate traits of POAG rather than the dichotomous disease status itself [37, 42, 68, 117, 126]. The use of these traits, or endophenotypes, has been well recognised in psychiatric research as a method of quantifying and deconstructing complex diagnoses [127]. Traits such as IOP, VCDR, ODA and CCT offer objective, quantitative data, and provide more statistical power for analyses than a subjective, dichotomous POAG diagnosis [42]. The use of these traits is becoming more popular, not only to simplify POAG into some of its components, but also to gain a better understanding of the biological pathways involved in POAG pathogenesis. Complex diseases, such as POAG may have several endophenotypes and these endophenotypes may also be linked to diseases other than POAG [37].

For a trait to be useful as an endophenotype for a disease, it must be heritable, be genetically correlated with the disease of interest, but not cause the disease on its own, and be able to be measured in both healthy and diseased individuals [37, 42]. A recent systematic review and meta-analysis of relevant POAG endophenotypes has shown the heritabilities to be high, with pooled estimates for; IOP 0.43 (0.38-0.48), CDR 0.56 (0.44-0.68) and CCT 0.81 (0.73-0.87) [128]. Several studies have specifically focused their POAG research on these quantitative traits, for example; [29, 93, 105, 110, 129-133]. Heritabilities accounted for by variants identified in GWAS specifically conducted on these endophenotypes are greater than those from POAG GWAS. The top SNPs from a GWAS using an optic cup measurement explained between 2.1% and 3.2% of the variance of POAG in the two independent cohorts of that study [134]. Additionally, a GWAS identifying novel IOP loci explained 17% and 9% of variance in the EPIC-Norfolk and UK Biobank cohorts respectively, with the latter estimate predicted to be more accurate due to decreased IOP measurement error, as multiple measurements were taken [135].

Iglesias et al. (2015) [37], have compiled a comprehensive list of the genes, which are the closest to GWAS hits, implicated in the POAG endophenotypes and presented them as Venn diagrams (Figure 1-6). These diagrams show that there are many genes which have only been implicated with individual traits and others which are involved with more than one trait. Although the ocular traits are under the control of many genes, the genes may not necessarily be involved with POAG. Understanding how they affect POAG pathogenesis will be simplified by knowing which quantitative trait they affect. A genetic correlation between the endophenotype and POAG is a prerequisite for identifying variants and genes important in POAG pathogenesis. Charlesworth

et al. (2010) [42], demonstrated that IOP and VCDR were genetically correlated with POAG in a study on 22 large Caucasian families. Other studies have demonstrated that variants statistically associated with IOP and VCDR were also significantly associated with POAG [66, 97, 131, 136].

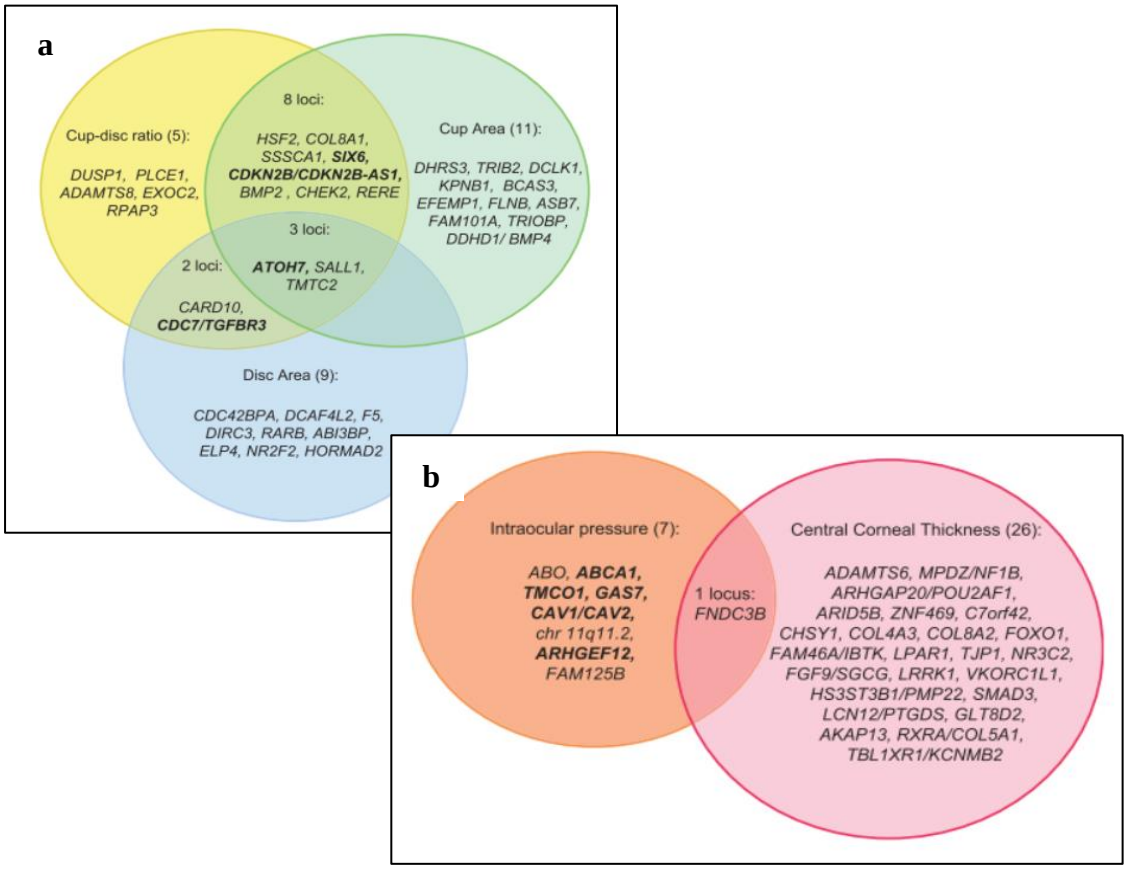


Figure 1-6 Genes implicated in POAG and its intermediate traits. The Venn diagrams illustrate the number of loci that show association with (a) optic disc parameters and (b) intraocular pressure and central corneal thickness. Genes associated with only one trait are listed below the trait name. Genes associated with more than one trait are listed in the overlapping regions of the diagram. Genes which have also been associated with POAG are listed in bold. Source: Iglesias et al. (2015) [37]

1.11 The use of next generation sequencing technologies to find disease causing genes

Next generation sequencing (NGS) technologies, encompassing whole genome sequencing (WGS) and whole exome sequencing (WES), are newer techniques which are being used to uncover the genetic causes of disease. With the huge reduction in costs of these technologies in recent years, they have become more accessible to researchers as additional methods to finding

disease causing mutations [137]. WGS and WES enable the full spectrum of variants to be identified, from common (MAF ≥ 5%) to uncommon (1% ≤ MAF < 5%), rare (MAF < 1%), and private (restricted to a family) variants [138, 139]. Using NGS technologies in family-based studies provides an exciting opportunity to search for these rare and private variants which may be important in disease [138, 140].

WES produces sequencing data for the 1% of the genome which is protein coding (~30Mb). Although most of the genome is not sequenced, it is believed that many disease causing mutations are within these coding regions [139]. So although it is recognised that not all disease causing mutations will be identified using WES alone, there is still the potential to find many important variants. The cheaper costs associated with WES, compared to WGS, allow for more study participants to be sequenced [137]. Although the information which is provided by NGS is detailed and valuable, this technology has its limitations. Both WGS and WES produce large amounts of data which need to be stored and managed. Specialised bioinformatics skills and tools are needed to handle the data and identify variants of interest [141, 142]. Potentially deleterious variants need to be identified from within the "noise" of sequencing errors and non-deleterious mutations. WES has additional limitations as structural variations, for example CNVs may not be able to be detected and not all exons are actually sequenced [139]. Although there are obvious limitations to NGS technologies, they are outweighed by quality of information revealed. NGS provides a more comprehensive genotyping tool than has previously been available, and can be used in family-based linkage analyses and population-based association studies.

In POAG research, recent studies using WES have confirmed previous findings and identified new variants in genes involved with the disease itself and with the intermediate, clinical traits of glaucoma [100, 107, 143-149]. The use of NGS with linkage studies has huge potential for accelerating the identification of both common and rare variants in disease. Families with many affected members, who are said to be "enriched" for a disease, provide improved statistical power for identifying rare variants, as these rare alleles have a 50% chance of being passed from a parent to each child [125]. Already, family studies using WES have identified rare variants in optic disc parameters of POAG patients [145-147]. New approaches, such as developing identity by descent (IBD) algorithms from NGS data, are being established to cope with the increased complexity of using NGS information in linkage studies. These are allowing for larger families with complex structures to be used [150-152].

1.12 Hypothesis

POAG is a heterogeneous disease with a complex genetic component. Although there is an abundance of literature on suspected POAG risk alleles as well as some rare causative genes, very little is known about how these genes affect the pathogenesis of the disease. Researching the intermediate traits of POAG may simplify the process of gene identification for this disease. Using WES data from large families enriched for POAG, to conduct linkage analyses on these intermediate traits, may prove to be a valuable method for finding variants influencing these traits. These families may harbour rare, deleterious variants as well as common variants. Identification of these variants has the potential to implicate genes and their pathways important in the pathogenesis of POAG. Ultimately, with a greater understanding of the genetics of POAG, an earlier diagnosis may be possible as well as better, targeted treatments.

The use of extended pedigrees enriched for POAG, to identify genetic variants associated with its heritable clinical traits, will provide an improved understanding of the genetic susceptibility to glaucoma.

1.13 Aims

Aim 1

To determine the contribution of published genetic variants associated with POAG and related traits, to trait variance in five extended POAG enriched families.

Specifically, this study will use a family-based association analysis on published genetic loci to identify variants associated with POAG and its clinical measurements in the families of this study. Three sets of variant data will be examined from the association analysis:

- SNPs reported in the literature as associated with POAG and its intermediate traits
- 20kb windows around the reported SNPs from the literature
- Linkage regions reported in the literature as associated with POAG and its intermediate traits.

Aim 2

To identify quantitative trait loci for intermediate POAG traits using linkage analysis in five extended pedigrees.

Specifically, identity by descent will be estimated from whole exome sequencing data and used to conduct variance components linkage analysis on five POAG enriched families to identify quantitative trait loci for IOP and VCDR.

Aim 3

To propose candidate genes which may be involved in IOP regulation in the families of this study.

Specifically, this study will use *in silico* tools to investigate genetic variants within the most significant peaks identified from the linkage analysis in Chapter 4, and propose genes which may be involved in IOP regulation.

Chapter 2
General Methods

2.1 Ethics approval and consent

As described previously for these families [71, 153, 154], ethics approval was obtained from the Oregon Health and Science University (OHSU) Institutional Review Board (Portland, OR), the Human Research Ethics Committees of the Royal Victorian Eye and Ear Hospital (Victoria, Australia), and the Human Research Ethics Committee (Tasmania) Network. This study was conducted in accordance with the tenets of the Declaration of Helsinki. Written, informed consent was obtained from all participants.

2.2 Families used in this study

The families used in this study are part of a long standing collaboration of the Oregon Glaucoma Genetics Study (OGGS) [71, 153] and the Glaucoma Inheritance Study in Tasmania (GIST) [154]. The OGGS families are from the north-western United States of America and the GIST families are from the southern Australian island state of Tasmania and all families are of European descent. Recruitment of these families in their respective glaucoma studies has been previously described in detail [42]. Five pedigrees were selected for this project based on contribution to linkage signals in pilot work conducted as part of a NIH funded project in 2006 (R01 EY010555). That pilot study, and subsequent research using these families [42], included extensive relationship testing, conducted by Jac Charlesworth, to ensure correctness of the pedigrees prior to the commencement of this study. Table 2-1 provides summary information about these families. Three families from the OGGS and two families from the GIST were used. These large families span five to seven generations and range in size from 48 individuals in family GTAS54 to 201 in family 98002. A total of 531 family members were included in this study, with clinical data obtained from over 300 participants. DNA samples, for whole exome sequencing, were obtained from 249 of the family members.

Table 2-1 Families participating in OGGS and GIST selected for this study. Number of family members with clinical and genetic data are shown.

Family name	Population	Total (N)	Generations in family *	Sex (F) %	Average age (years)	Clinical data (N)			WES (N)
						IOP	VCDR	CCT	
93001	OGGS	138	5	49.3	43.2 ± 21.0	84	85	66	57
95002	OGGS	60	5	51.7	56.8 ± 17.4	40	41	28	36
98002	OGGS	201	7	51.7	59.5 ± 17.2	107	104	70	91
GTAS04	GIST	84	6	53.5	59.9 ± 15.4	41	42	9	37
GTAS54	GIST	48	6	52.1	60.9 ± 13.7	29	29	13	28

*** =** informative generations with clinical and/or DNA data
OGGS = Oregon glaucoma genetics study, **GIST** = glaucoma inheritance study in Tasmania
Average age is ± standard deviation.
IOP = intraocular pressure, **VCDR** = vertical cup to disc ratio, **CCT** = central corneal thickness
WES = whole exome sequencing

Detailed pedigree diagrams of the five families are shown in Figures 2-1 to 2-5 on the following pages, indicating POAG disease status as well as the clinical measurements for intraocular pressure (IOP), vertical cup to disc ratio (VCDR) and central corneal thickness (CCT). Individuals with a medication adjusted IOP (as described below in section 2.5.1) are also indicated in these pedigrees. Families 93001, 95002 and 98002 are from the OGGS in America and families GTAS04 and GTAS54 are from the GIST in Australia. These families all represent the adult onset form of POAG, with diagnosis usually after the age of 40 years for those diagnosed. Family members indicated with no POAG status in the pedigree diagrams (Figures 2-1 to 2-5) were either not clinically assessed for POAG or were too young at their final clinical examination to definitively categorise them with a negative diagnosis for POAG. Although the POAG status of each individual is shown in the following pedigrees, our study did not use POAG as the phenotype of interest. The IOP, VCDR and CCT measurements were used for all of the analyses conducted in this study, with POAG status only revealed at the end of the study, when variants of interest were being analysed.

Family 93001

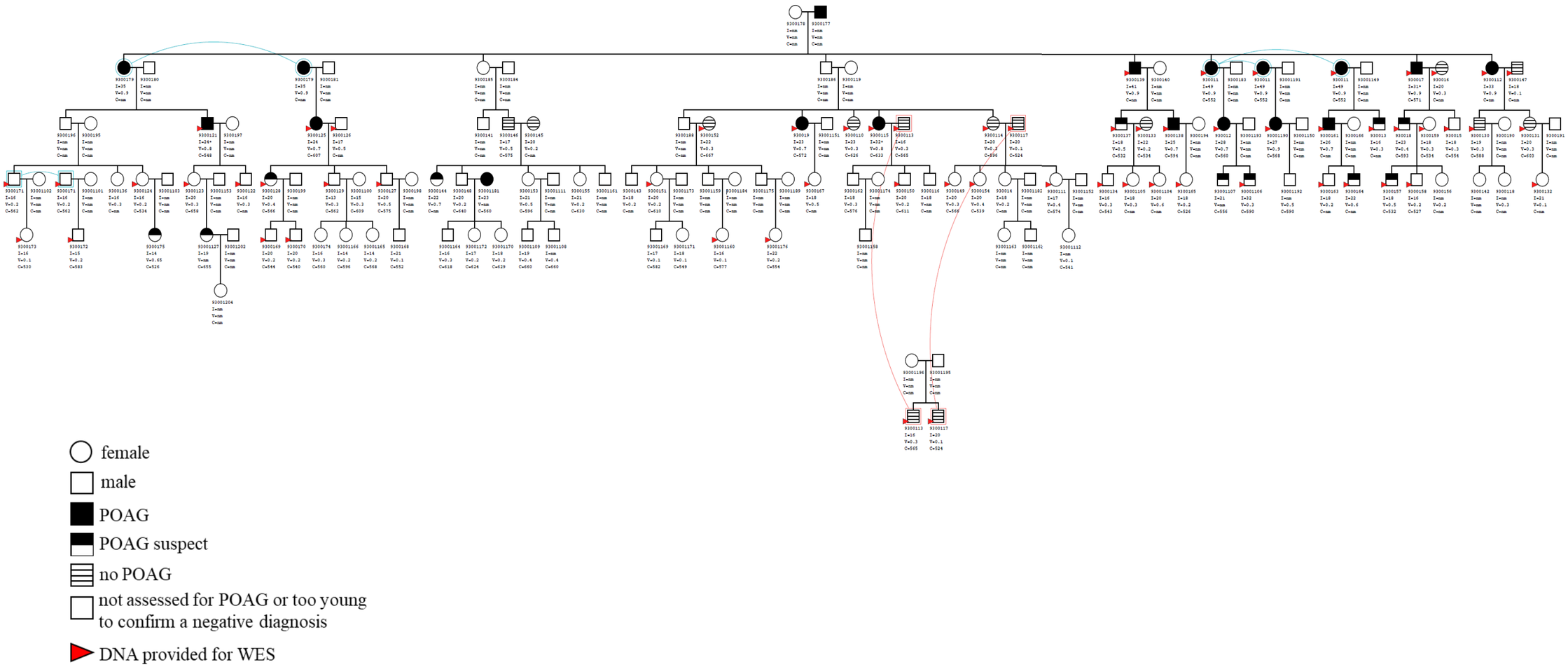


Figure 2-1 **Family 93001** Identification number = the first string of numbers, **I** = IOP measurement (mmHg), ***** = medication adjusted IOP, **V** = vertical cup to disc ratio, **C** = CCT measurement (µm), **nm** = not measured, coloured lines represent the same individual appearing more than once in the pedigree

Family 95002

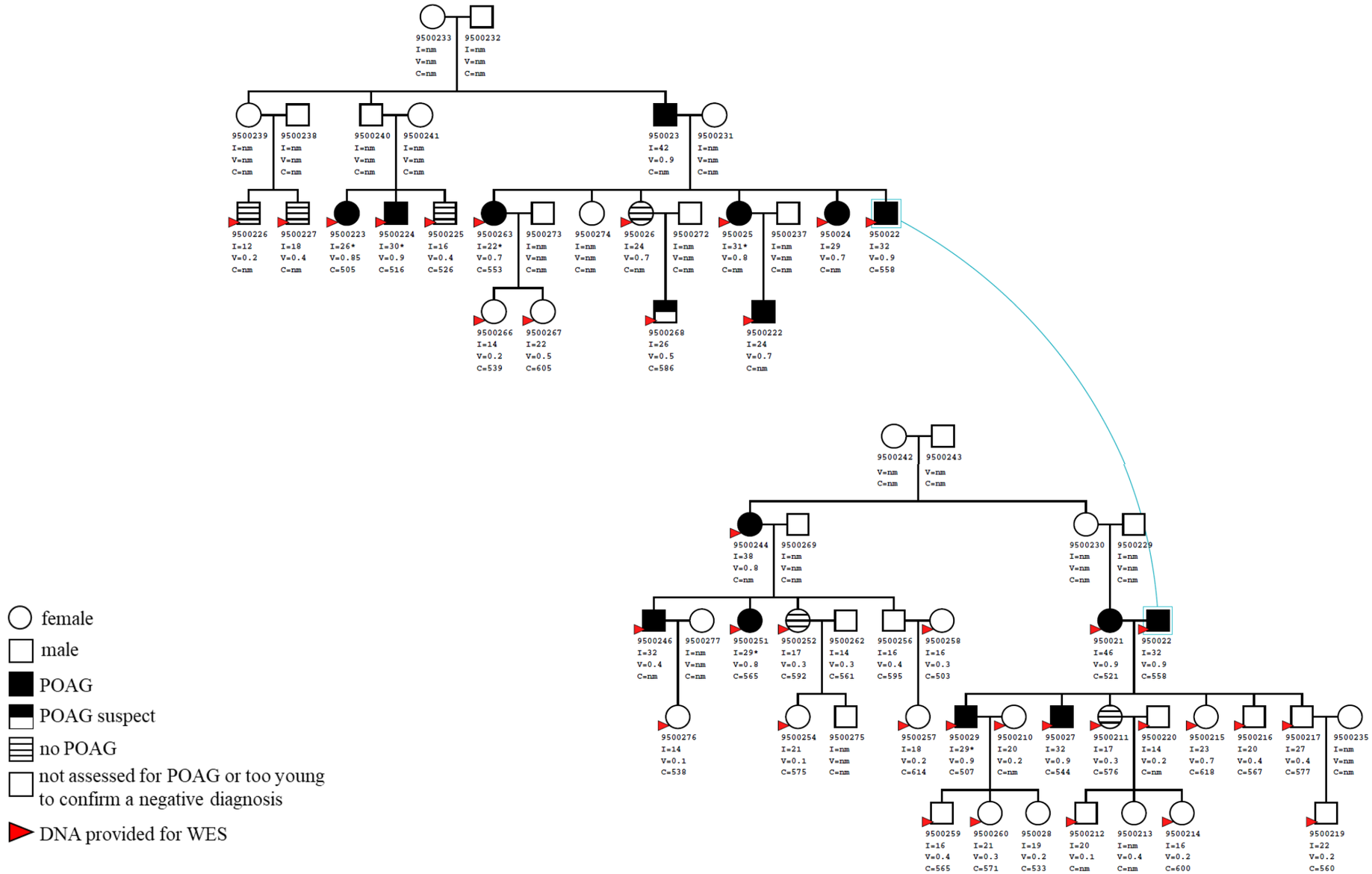


Figure 2-2 **Family 95002** Identification number = the first string of numbers, **I** = IOP measurement (mmHg), ***** = medication adjusted IOP, **V** = vertical cup to disc ratio, **C** = CCT measurement (µm), **nm** = not measured, coloured lines represent the same individual appearing more than once in the pedigree

Family 98002

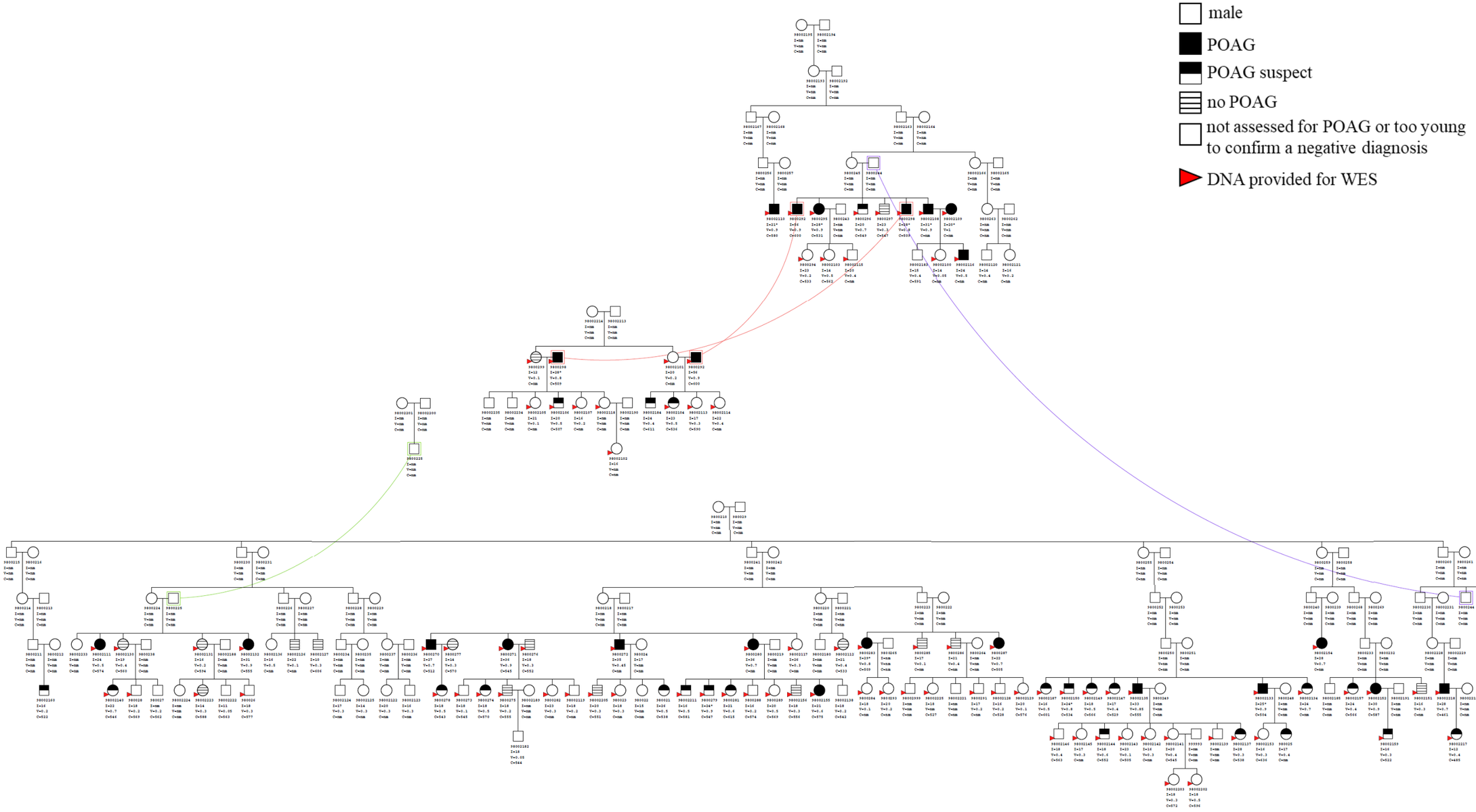


Figure 2-3 **Family 98002** Identification number = the first string of numbers, **I** = IOP measurement (mmHg), ***** = medication adjusted IOP, **V** = vertical cup to disc ratio, **C** = CCT measurement (µm), **nm** = not measured, coloured lines represent the same individual appearing more than once in the pedigree

Family GTAS04

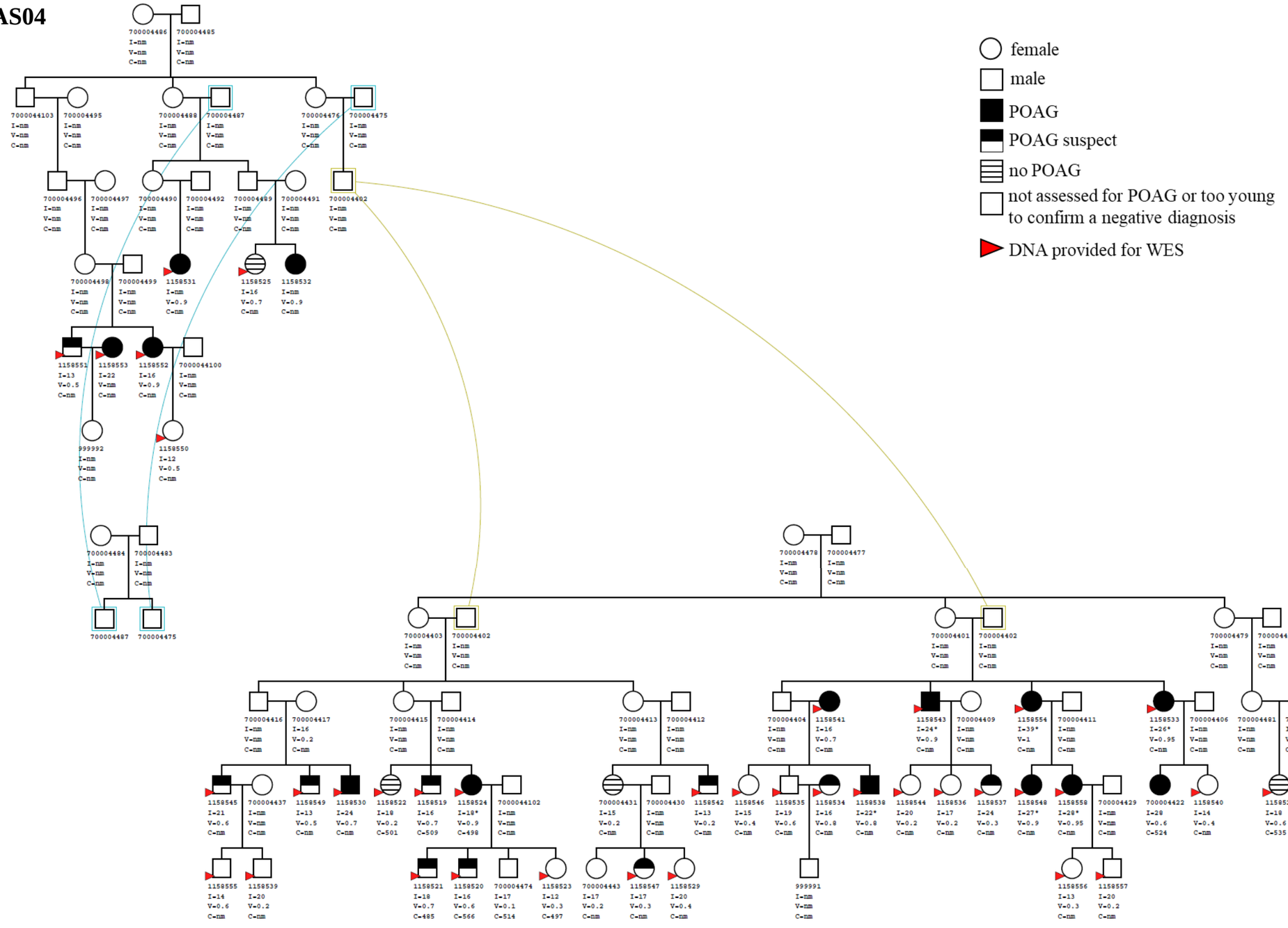


Figure 2-4 Family GTAS04 Identification number = the first string of numbers, **I** = IOP measurement (mmHg), ***** = medication adjusted IOP, **V** = vertical cup to disc ratio, **C** = CCT measurement (µm), **nm** = not measured, coloured lines represent the same individual appearing more than once in the pedigree

Family GTAS54

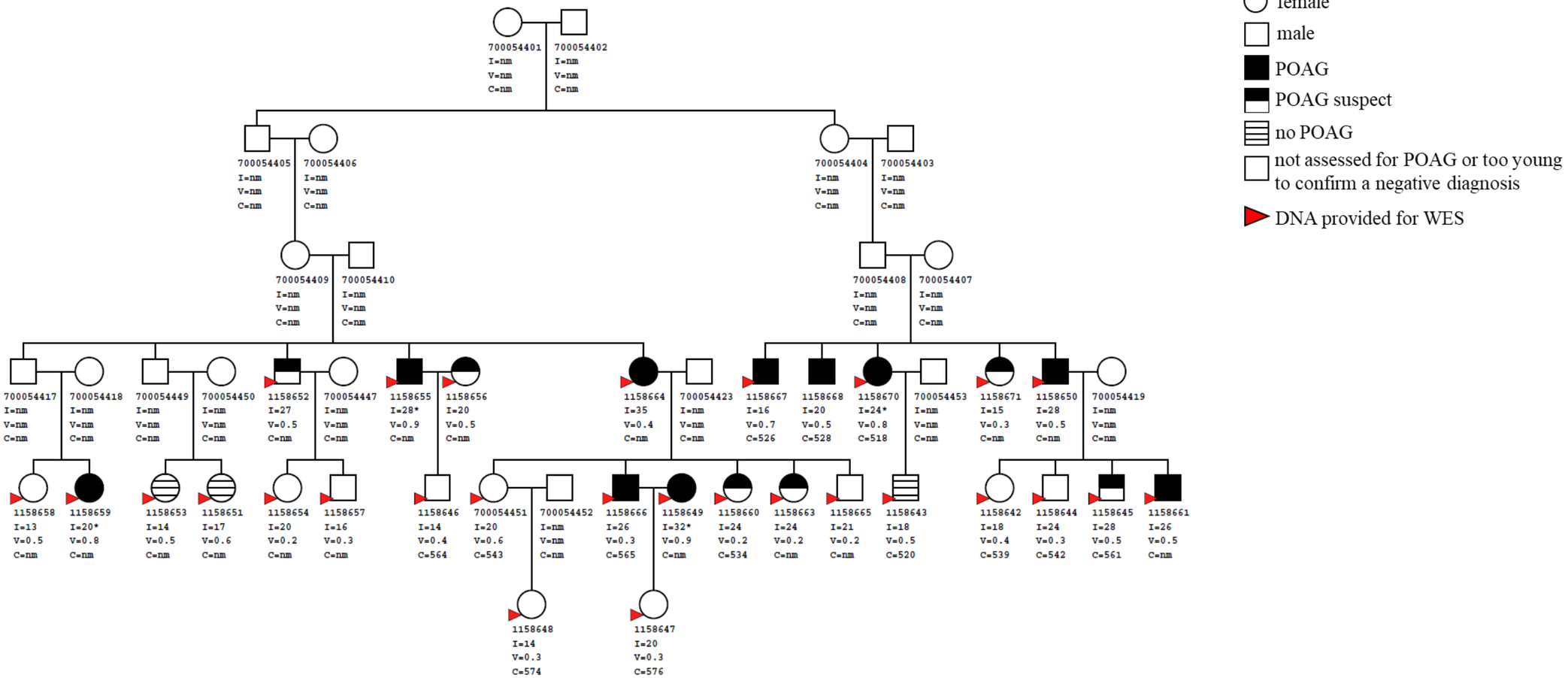


Figure 2-5 Family GTAS54 Identification number = the first string of numbers, **I** = IOP measurement (mmHg), ***** = medication adjusted IOP, **V** = vertical cup to disc ratio, **C** = CCT measurement (μm), **nm** = not measured

2.2.1 Family 93001

Family 93001 (Figure 2-1) is a 5 generation family consisting of 138 members. The original founders of the family had 7 children, 5 of whom were diagnosed with POAG. In total, 15 family members were diagnosed with POAG, with a further 11 diagnosed as suspected POAG cases at their last clinical examination. Most of the POAG and suspected POAG cases are in the 2nd and 3rd generations of the family, with the later generations possibly too young to develop POAG by their last examination. Clinical data and DNA samples were obtained from most of the direct descendants of those diagnosed with POAG. In total 85 family members provided clinical data and 57 provided DNA for sequencing.

2.2.2 Family 95002

Family 95002 (Figure 2-2) is the smallest of the American families, with 60 members. This family was not descended from the same founder, but were linked together by a couple (ID:950021 and ID:950022), both of whom were from POAG enriched families. In total, 14 family members were diagnosed with POAG plus one suspected POAG case. Two thirds of this family provided clinical data and 36 members provided DNA samples.

2.2.3 Family 98002

Family 98002 (Figure 2-3) is the largest of the five families involved in this study, with 201 individuals included. There is one main family, descended from founders ID:9800210 and ID:980029. A marriage loop in the 4th generation between two brother and sister pairs is indicated in the smaller pedigree diagrams. The POAG cases are spread across this whole family with clinical measurements and DNA samples obtained from most of individuals in the 4th to 7th generations.

2.2.4 Family GTAS04

Family GTAS04 (Figure 2-4) is the larger of the two Tasmanian families and demonstrates a more complex family structure. The founders in the main part of the pedigree, ID:700004478 and ID:70000477, had three daughters. Two of these daughters had families with the same partner, ID:700004402, and there are POAG affected descendants in both families. The partner's side of the family is linked to the main pedigree and shown in the upper left of Figure

2-4, with three additional POAG and one suspected POAG case. Clinical data and DNA samples were obtained from most of the individuals with at least suspected POAG, their siblings and descendants.

2.2.5 Family GTAS54

GTAS54 (Figure 2-5) is the smallest of the five families included in this study, with 48 individuals, and has a single lineage group. The generations enriched for POAG are the great grandchildren of the original founders, ID:700054401 and ID:70005402. In total, 16 were diagnosed with at least suspected POAG (included one spouse). Nearly all of the great-grandchildren of the founders and their descendants provided both clinical data and DNA samples for this study.

2.3 Clinical data

Clinical data was collected from participants for POAG disease status and for the POAG endophenotypes; intraocular pressure (IOP), vertical cup to disc ratio (VCDR) and central corneal thickness (CCT) and has been described in detail elsewhere [42, 155]. Briefly, the clinical assessment for glaucoma disease status involved assessing the appearance of the optic nerve head in conjunction with visual field defects typical of glaucoma. Examination included measuring the level of optic nerve head excavation accompanied by thinning of the neuroretinal rim, which may also show notching, pitting, Drance-type nerve fibre layer haemorrhages or a general loss of retinal fibre layer. For a diagnosis of POAG, the iridocorneal angle was open and there was no indication of other causes for the glaucoma, such as trauma, medication use or pseudoexfoliation syndrome. The classification of individuals with a definitive versus suspected POAG diagnosis is detailed elsewhere [153, 155], and is based on the degree of optic nerve cupping and visual field loss.

IOP was measured with Goldmann applanation tonometry with multiple measurements taken over several years for most individuals. The maximum recorded IOP was used as the trait of interest in this study. Medical records were accessed for those individuals with a POAG diagnosis prior to the study, to obtain any pre-medication IOP data. IOP was measured in 301 family members. The VCDR, as a measure of optic nerve head excavation and damage, was classified using a slit lamp biomicroscope, by one or two clinicians during the clinical examination process. Additionally, stereo fundus photographs were taken for future reference

and assessed by a third clinician if there was a discrepancy between the examining clinicians. The larger VCDR of the two eyes, for each participant, was used as the trait for this study, with the maximum measurement recorded used as the trait in this study. VCDR was measured in 301 family members. CCT was measured with ultrasound pachymetry, with 25 measurements taken for each eye at each timepoint. The mean of these measurements was recorded at each timepoint and the overall minimum record used as the trait for this study. CCT was measured in 186 family members. VCDR and CCT were also measured over multiple time-points.

2.4 Whole exome sequencing data preparation

2.4.1 Sequencing, alignment and variant calling

The DNA sequencing was conducted in the laboratory of our close collaborator, Professor Joanne Curran, at the Texas Biomedical Research Institute, USA. DNA samples from 249 members of five families were extracted from whole blood and prepared for sequencing using the Nextera Expanded Exome Enrichment Kit (Illumina: https://sapac.illumina.com/products/by-type/sequencing-kits/library-prep-kits/truseq-exome.html). 100bp paired-end sequencing was conducted on the Illumina HiSeq 2000 platform. The expanded exome kit provided 62Mb exome content including 201,121 exons in 20,794 genes, as well as UTRs and miRNA sites. The Churchill pipeline [156] was used for sequence alignment. Within this pipeline, BWA-MEM (version 0.7.12) [157] was used to align fastq sequences to the hg19 human reference genome, Picard (version 1.114) (https://broadinstitute.github.io/picard/) was used for deduplication, and the Genome Analysis Toolkit (version 3.6) (GATK, https://software.broadinstitute.org/gatk/) [158] for local realignment around indels and recalibration of base quality scores. The BAM files which were generated were the starting point for this research project. A bcbio pipeline (https://github.com/bcbio/bcbio-nextgen) was used to call SNPs and indels from the BAM files. The configuration file used for this variant calling is provided in Appendix A.1. Within this pipeline, variants were joint called using GATK (version 4.0.3.0) [159].

2.4.2 Annotating the VCF file

After variant calling, the resultant VCF file was labelled with quality tags using the GATK "Variant Filtration" tool (see Appendix A.2.1). Variants with a depth of coverage (number of

filtered reads) of 10 or greater were annotated as “high-coverage”. Variants with a Phred-scaled genotype quality score of 20 or greater were annotated as “high-quality”. “High-confidence” calls were both high-coverage and high-quality. Conversely, “low-confidence” calls were both low-coverage and low-quality. Genotypes could also be assigned low-coverage with high-quality and high-coverage with low-quality. These quality annotations were used in downstream analysis of variants of interest.

ANNOVAR (April 2018 version) [160, 161] was then used to annotate the vcf file for specific gene information, population frequencies in multiple populations, predicted scores of deleteriousness and phylogenetic conservation scores. The ANNOVAR command is shown in Appendix A.2.2.

2.5 Traits used in this study

2.5.1 IOP and IOPMed

Both an unadjusted maximum IOP measurement (IOP) and a medication adjusted maximum IOP measurement (IOPmed) were used in the course of this study. Where possible, IOP measurements recorded prior to treatment were used. For those individuals who had commenced pressure reducing medication prior to this study, a correction factor was applied to the IOP measurement, similar to that applied by Springelkamp et al. (2017) [136]. The standard deviation (SD) of the maximum IOP measurements from all 301 participants was calculated (SD = 6.1mmHg) and a 1 x SD was added to the IOP measurements for the 29 individuals who only had post-medication measurements. This trait was called IOPmed, to distinguish it from the unadjusted IOP trait.

2.5.2 VCDR and CCT

The VCDR and CCT measurements (as described in section 2.3) were used without any form of medication adjustment in this study. The raw VCDR measurements were a ratio and ranged between 0 and 1. CCT measurements were used in μm.

2.6 Variance components analysis of traits

The Sequential Oligogenic Linkage Analysis Routines (SOLAR) package (version 8.8.1) [162] was used to conduct the quantitative genetic and linkage analyses for this project. SOLAR uses variance components statistics, which is discussed in detail in Chapter 4. SOLAR was selected based on the size of the pedigrees and the computational limits of most of the other packages for pedigrees of this size.

2.6.1 Variance components modelling of traits

Variance components modelling was conducted using SOLAR to determine the heritability of the traits as well as the contribution of the covariates to the phenotypes. Covariates included were; age, sex, the interaction of age and sex, age^2, and the interaction of age^2 and sex. The `polygenic -screen` command calculated the significance of each of the covariates and those not reaching a nominal correction significance of $p < 0.1$ were removed from the final polygenic model. Modelling was conducted separately for the full pedigree file, consisting of all 531 family members, as well as an adjusted pedigree file, consisting of only the 249 family members with WES data. The full pedigree file was used for the measured genotype association analysis (MGA) analysis, discussed in section 2.6.3 below. The adjusted pedigree file was used for the linkage analysis and is discussed fully in Chapter 4.2.1.3.

2.6.2 Data transformation

All traits were normalised prior to the MGA (discussed below). Variance components polygenic modelling was performed with all covariates included, unlike the general modelling described above. These covariates were included in the model regardless of their significance through the `polygenic -all -screen` command. Residuals for the maximised model were computed for each trait. An inverse normal transformation of the residuals was computed using the `inormal` command. The covariate adjusted, inverse normalised residuals were then treated as a trait and used in the MGA analyses described below.

The raw VCDR measurements, which range between 0 and 1, were multiplied by 10 to increase the standard deviation of the data, reducing errors in the analysis pipeline. After analysis, all results for VCDR were converted back to the true ratios.

2.6.3 Measured genotype association (MGA) analysis

MGA is a test to determine whether the phenotypic means of a trait are affected by genotypic differences [163, 164]. Trait means are estimated for each of the genotypes, with an additive model used which constrains the heterozygote to be half way between the two homozygotes. SNPs are included individually as covariates in the model, scored as a dosage of the minor allele. A likelihood ratio test is used to test if the regression co-efficient of each SNP genotype differs from zero. Conducting MGA within a variance components framework allows for the non-independence of family relationships to be taken into account, which would otherwise inflate the p-values [165-167]. MGA uses both genotypic and phenotypic data for each family. Any individuals without phenotypic data were not included in the analysis.

As MGA uses the number of rare alleles, the genotypes in the full VCF file were converted to a dosage file representing the number of rare alleles. SNPs and Indels were handled separately for this process and subsequent analysis. A custom R script (see Appendix A.3) was used for this conversion, as well as running quality control on the data. Variants were filtered based on the rarity of the minor allele, with a minor allele copy (MAC) number of less than 5 removed. The quality control step removed variants with at least 50% of missing calls. Table 2-3 summarises the number of SNPs and indels at each of the major stages of the dosage file preparation.

Table 2-2 SNPs and indels remaining at the major stages of dosage file preparation

	No. SNPs	No. Indels
Start	429,670	55,499
After MAC filtering	151,253	25,185
Final (after quality control)	142,210	22,002

MAC = minor allele copy

The completed dosage file was then used to run the MGA analysis in SOLAR. MGA was run separately for each of the three clinical traits; IOP, VCDR and CCT. The commands used for this analysis are included in Appendix A.4. The MGA used the full pedigree file listing all 531 family members and their relationships as well as the full phenotype file. Based on this pedigree information, SOLAR was used to calculate a kinship matrix which was used to identify the family relationships which were necessary for accurate MGA analysis. MGA was run twice for each trait; once using the normalised phenotype values (as described above in section 2.6.2) and once on the raw, non-normalised phenotype values. Normalising the traits provides

accurate p-values for the association analysis but is not appropriate for determining the magnitude of the change in phenotype, due to the transformation of the data. The MGA analysis on the raw phenotype values was used to provide beta-values indicating the effect size of each allele.

2.7 Additional information

2.7.1 R scripts

R (version 3.3.1) [168] was used to compose most of the custom data manipulation scripts for this study. The R package, ggplot2 [169], was used to draw the linkage diagrams in Chapters 4 and 5. An example linkage diagram script is provided in Appendix A.5.

2.7.2 Pedigree drawings

Pedigree diagrams were drawn using Cranefoot (version 3.2.3) [170]

Chapter 3
Contribution of reported POAG associated loci with IOP, VCDR and CCT in five extended families

3.1 Introduction

To date there have been hundreds of loci associated with POAG risk; ranging from individual SNPs and small indels, to whole genes through to large linkage regions encompassing many genes. Some of the loci are well replicated and recognised as true POAG associated loci, for example; *TMCO1* [12, 66, 86, 89, 91, 95, 109, 136, 171], *CDKN2B-AS1* [12, 29, 69, 91, 92, 94, 111] and *SIX1/SIX6* [29, 69, 89, 90, 94, 95, 97, 111, 136], although the function of these loci in POAG pathogenesis is still not well understood. Other POAG loci have not been replicated in independent studies and their involvement with POAG are not necessarily confirmed, for example; *UST* [172], *FREM3* [143] and *DERA* [172]. Linkage regions, often encompassing millions of base pairs, may be family specific, for example; the GLC1D [173], GLC1F [71] GLC1L [174] and GLC1P [61] loci. Other linkage regions were identified using several families, for example; the GLC1B [175], GLC1H [176] and GLC1I [177] loci. Linkage studies have successfully been used as the first stage in identifying regions of the genome of interest, allowing for candidate gene studies to follow. *MYOC*, *OPTN*, *WDR36* and *CYP1B1* have successfully been identified as glaucoma causing genes where linkage studies have identified the initial genomic location [46, 50, 70, 73]. However, linkage studies are difficult to replicate as the families ascertained for these studies may differ in their size, structure, disease phenotype and ethnicity.

A recognition of the importance identifying loci involved with the clinical intermediate traits of disease, rather than the discrete glaucoma disease diagnosis itself, has generated research using a variety of strategies. GWAS, candidate gene studies and linkage studies have been conducted focussing on IOP, VCDR, CCT and other optic disc measurements, which are involved with the POAG disease process either directly, or as a risk factor for the disease [29, 93, 105, 110, 129-133]. IOP and VCDR have already been identified as ideal endophenotypes for POAG in the families of this study [42]. Although CCT was not genetically correlated with POAG in these families [42], it is still a recognised risk factor for the disease [2, 13, 178]. As

a highly heritable trait, results may still indicate important genes involved in CCT biology, although they may not be involved with POAG pathogenesis in these particular families.

To determine whether POAG loci from the literature are associated with clinical measures in the families of this study, firstly, a comprehensive literature review was undertaken, encompassing relevant loci from a variety of study types; GWAS, candidate gene studies and linkage studies. A family-based association test (measured genotype association; MGA) was conducted on these identified loci for each of the intermediate clinical traits measured in the families. Association was tested with three loci types; with the published SNP loci, in 20kb windows around each SNP, to capture variants in LD with those published, and across POAG linkage regions, which encompass much larger regions of the genome. The purpose was not to directly compare results from this study with the hundreds of published POAG papers, but to ascertain whether the families of this study share important glaucoma loci with those already identified.

3.1.1 Aim of this study

To determine the contribution of published genetic variants associated with POAG and related traits, to trait variance in five extended POAG enriched families.

Specifically, this study will use a family-based association analysis on published genetic loci to identify variants associated with POAG and its clinical measurements in the families of this study. Three sets of variant data will be examined from the association analysis:

- SNPs reported in the literature as associated with POAG and its intermediate traits
- 20kb windows around the reported SNPs from the literature
- Linkage regions reported in the literature as associated with POAG and its intermediate traits.

3.2 Methods

Figure 3-1 outlines the stages involved with conducting a comprehensive literature review of POAG loci through to running the family-based association analyses on these identified regions.

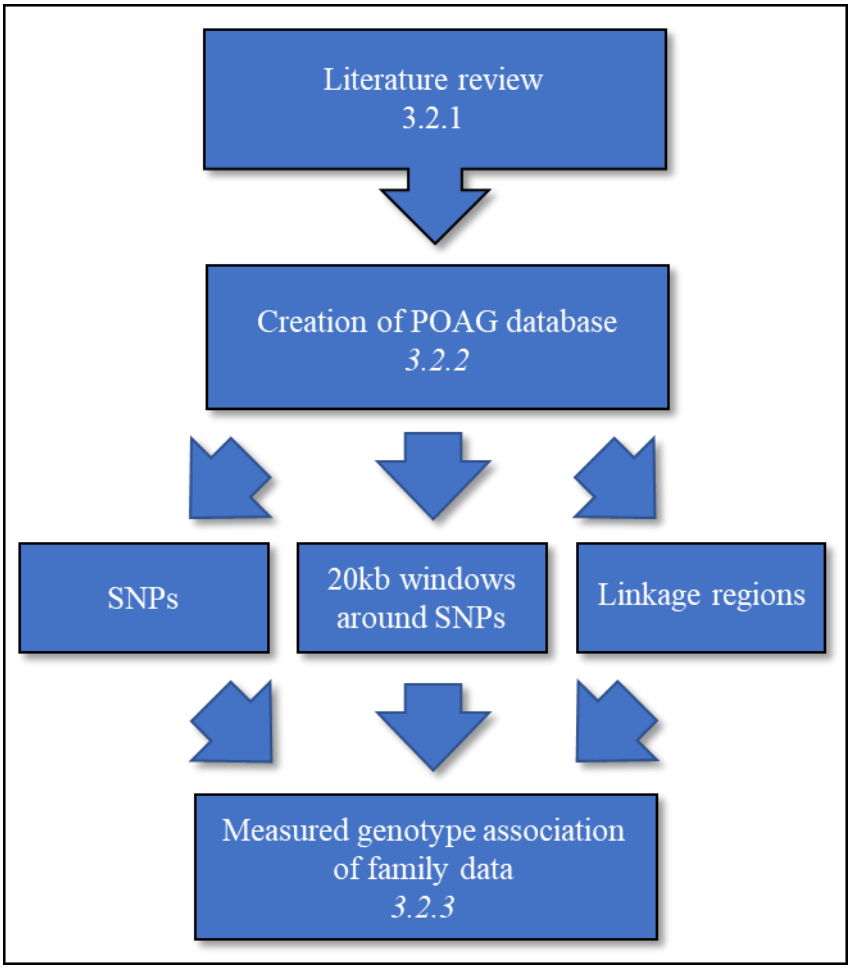


Figure 3-1 Major steps involved with conducting measured genotype association analysis from a literature review of POAG loci. Italicised numbers refer to the following sections with detailed descriptions of each step.

3.2.1 Literature review

A literature review, to gather information on published POAG loci, was initiated in 2016 and concluded in 2018, and included papers published between 1993 and 2018. A PubMed search was undertaken, using the following search terms:

```
(((((("Glaucoma, Open-Angle/genetics"[Mesh]) OR "POAG"[Title/Abstract]) OR
"primary open angle glaucoma"[Title/Abstract])) AND (((((("Genetics"[Mesh])
OR gene*[Title/Abstract]) OR locus[Title/Abstract]) OR
loci[Title/Abstract]) OR SNP[Title/Abstract]))) AND "Humans"[Mesh]) AND
English[Language]
```

Additional PubMed searches were also undertaken to ensure a comprehensive literature search.

```
((endophenotype) OR intermediate phenotype) AND glaucoma
(open angle glaucoma) AND genome wide association stud*
```

In addition to these searches, other publications were identified from the bibliography sections of papers already obtained. Several review articles and meta-analyses were used to ensure inclusion of as many relevant loci as possible and these provided many of the loci incorporated into the database [13, 27-29, 103, 111, 134, 136, 171, 179].

Detailed results and information on study participants was extracted from each paper. Studies involving participants of European or Caucasian descent were prioritised, to obtain as many loci as possible applicable to the families of European descent of this current study. Studies involving non-European populations were included if they were a part of larger studies, but smaller non-European studies were not prioritised in this literature review.

3.2.2 Creation of POAG loci database

Information extracted from the literature review was compiled into a database (Appendix B). Loci associated with the POAG clinical status were included as well as the IOP, VCDR and CCT clinical traits. Some studies reported associations with optic disc area (ODA) and optic cup area (OCA) which, although not included as endophenotypes in this study, are relevant traits to assess [29, 180]. The most significant results as reported by the authors were included, even if they did not necessarily reach genome-wide significance after Bonferroni correction, to ensure important findings were not omitted due to a subjective definition of significance. Some variants, for example rs7518099 in *TMCO1* and rs10483727 at the *SIX1/SIX6* locus,

demonstrated association in some studies but not others [29, 91, 93]. Variants were included if they were associated with POAG or one of the clinical traits in at least one study.

The loci included in the database in Appendix B are a comprehensive representation of the major POAG and endophenotype loci published in the literature, especially for European participants. Major glaucoma linkage regions, designated GLC1A-Q were also included, as well as other linkage regions without designated names. The database provides the most accurate location of each of the loci as possible, from the original publications. Some papers reported microsatellite markers or particular regions of genes (for example the promoter region of *NPPA* [181]) and these were converted to the most accurate base pair positions in the hg19 reference genome.

Base pair position and POAG and/or clinical trait associations from the literature were extracted from Appendix B and used to test for association in the families. Association in the families was tested for variants falling into each of three groups from the literature; SNPs, 20kb windows around these SNPs, and linkage regions. Methods for each category varied slightly, as outlined below.

3.2.2.1 SNPs from the literature

A single list containing each variant's chromosome number and base pair position was generated from Appendix B. There were 448 SNPs identified in the POAG database. The VLOOKUP function within Microsoft Excel was used to determine which of these SNPs were captured in the whole exome sequencing conducted on the families (as shown in the VCF file described in section 2.4.2). The SNPs identified in the family data were tested for association separately for the IOP, VCDR and CCT clinical traits.

3.2.2.2 20kb windows around SNPs from the literature

To test for association in the 20kb windows around the published SNPs, firstly the loci were categorised by their clinical associations in the literature. SNPs were separated into POAG, IOP, VCDR, CCT and the combined ODA or OCA groups, as identified by "1" in these columns in Appendix B. A BED file, consisting of the chromosome number and a column each for the start and stop base pair positions, representing a ±10kb region around each of the SNPs, was generated for each clinical trait. Regions were common between BED files if SNPs were associated with more than one clinical trait in the literature. Table 3-1 summarises the numbers of SNPs represented in each of the 5 BED files.

Table 3-1 Number of SNPs with reported associations represented in BED files for POAG and its clinical traits

POAG *	IOP	VCDR	CCT	ODA or OCA
191	108	107	96	110

POAG* = primary open angle glaucoma, including 13 SNPs identified as associated with normal tension glaucoma and 4 loci identified as associated with juvenile open angle glaucoma **IOP** = intraocular pressure, **VCDR** = vertical cup to disc ratio, **CCT** = central corneal thickness, **ODA** = optic disc area, **OCA** = optic cup area

A custom R script (see Appendix C.1.1) was used with each of the BED files to extract variants within each region from the family sequencing data. Regions reported in the literature as associated with POAG were assessed for association with each of the three clinical traits; IOP, VCDR and CCT. Regions reported for each trait were assessed against the same trait in the families with the addition of assessing IOP loci for association with CCT and vice versa, due to relationship between CCT thickness and IOP measurement (see Chapter 1.7). In addition, published ODA and OCA loci were tested for association with VCDR in the families as these traits are also related. Table 3-2 shows the analysis plan for examining the family-based association data within the 20kb regions around the SNPs.

Table 3-2 Analysis plan for assessing reported associations against the clinical traits measured in the families

		Family MGA data		
		IOP	VCDR	CCT
Regions with reported associations for:	POAG	✓	✓	✓
	IOP	✓		✓
	VCDR		✓	
	CCT	✓		✓
	ODA or OCA		✓	

Ticks indicate reported associations assessed against family MGA data for each of the clinical traits.
POAG* = primary open angle glaucoma, including 13 loci identified as associated with normal tension glaucoma and 4 loci identified as associated with juvenile open angle glaucoma **IOP** = intraocular pressure, **VCDR** = vertical cup to disc ratio, **CCT** = central corneal thickness, **ODA** = optic disc area, **OCA** = optic cup area

3.2.2.3 Variants within larger loci and linkage regions

One BED file was generated for all the larger loci (such as regions of genes) and linkage regions, consisting of the chromosome number, start and stop base pair positions as reported in Appendix B. No additional windows were added to these regions as they were already broad. The same custom R script as used in the 20kb windows discussed above (Appendix C.1.1) was

used to extract these 38 larger regions from the family sequencing data. MGA analysis was then conducted on these regions for each of the clinical traits; IOP, VCDR and CCT.

3.2.3 Family-based association testing of published POAG loci

MGA is a family-based association test conducted within a variance components framework (as described in Chapter 2.6.3). MGA analysis was completed separately for IOP, VCDR and CCT, providing association data for each variant within the regions identified from the literature. Variance components polygenic modelling was conducted on the IOP, VCDR and CCT quantitative traits in SOLAR (as described in section 2.6.3) using the full pedigrees to calculate empirical kinships. MGA was run for both the normalised and raw trait values, to provide accurate association statistics (normalised) as well as effect sizes of each variant (untransformed).

As whole exome sequencing (WES) data was obtained from the families of this study, providing the genotyping information for this analysis, only regions captured by the exome sequencing were able to be included in the MGA analysis. Reported loci in non-coding and intergenic regions were not captured

3.2.3.1 Significance levels for association testing

As the association analyses were tested on three different types of loci from the literature; SNPs, 20kb windows around SNPs and linkage regions, determining statistical significance was calculated accordingly for each.

The SNPs were tested for association with each of the clinical traits measured in these families. A Bonferroni correction was used to adjust for the number of independent loci which were captured in the family sequencing data as well as the three clinical traits tested with MGA. The 20kb windows around the SNPs were only tested for association with the traits as outlined in Table 3-2. The linkage regions were tested for association with IOP, VCDR and CCT and were combined with the POAG 20kb window regions, as these too were assessed for association with each of the three clinical traits. We hypothesised that there would be only one causative variant per region and therefore used the number of regions assessed for each trait to correct for the number of tests performed. The p-value used for significance used a correction accounting for both the number of loci represented and the number of clinical traits analysed

for each list of regions (Table 3-2). The p-value used for suggestive significance only corrected for the number of loci represented but not the number of traits analysed.

3.2.3.2 Combine MGA data with annotated VCF files

To contribute additional information on the variants tested for association, annotations for each of the variants were added to the MGA output files. The variants in these regions were annotated using ANNOVAR (as described in section 2.4.2), providing information such as allele frequencies in different populations and measures of deleteriousness. Variants were also annotated for call quality, using GATK (as described in section 2.4.2). Only variants with at least 80% high confidence calls were used for further analysis. A custom R script (Appendix C.1.2) was used to join the MGA results to the extracted and annotated VCF files.

3.3 Results

3.3.1 Polygenic modelling of IOP, VCDR and CCT

Polygenic modelling for each of the clinical traits is displayed in Table 3-3. The three traits were all significantly heritable, with p-values ranging from 1.80 x $10^{-}9$ to 9.00 x 10^{-7}. CCT was the most heritable trait at 84% and the normalised VCDR trait least heritable at 40%. Only the non-normalised IOP trait was kurtotic and normalising the trait increased its heritability to 50%. Age was a significant covariate for each of the clinical traits. Covariates contributed most to the variance of VCDR, with 34%, and least to CCT, with 6%.

Table 3-3 Variance components polygenic modelling of normalised and non-normalised POAG clinical traits

Trait	h^2r ± SE (%)	p-value	Residual kurtosis	Covariates included in final model (non-normalised traits only)	Proportion of variance due to covariates (%)
IOP	44 ± 10	1.83 x 10^{-8}	**3.97**	age, age^2	18
IOP_norm	50 ± 10	1.36 x 10^{-8}	-0.19		
VCDR	43 ± 11	9.00 x 10^{-7}	-0.38	age, age^2	31
VCDR_norm	40 ± 11	5.00 x 10^{-7}	-0.19		
CCT	88 ± 12	1.80 x 10^{-9}	0.37	age	6
CCT_norm	84 ± 14	2.20 x 10^{-9}	-0.24		

POAG = primary open-angle glaucoma, **IOP** = intraocular pressure, **VCDR** = vertical cup to disc ratio, **CCT** = central corneal thickness, **norm** = normalised trait,
h^2r = heritability, **SE** = standard error, **residual kurtosis** = a measure of non-normal trait distribution, bold indicates kurtosis exceeding 0.8 threshold

3.3.2 Significance thresholds for each analysis

As described in section 3.2.3.1, significance levels were based on the number of loci captured in the families' sequencing data. With 41 SNPs from the literature returned in the family-based association data, Bonferroni corrected p-values of 4.1 x 10^{-4} (0.05 ÷ (41 SNPs x 3 traits)) and 0.0012 (0.05 ÷ 41 SNPs) were used for significance and suggestive significance respectively.

For the 20kb windows around the published SNPs, levels of significance were based on the combinations identified in the analysis plan in Table 3-2 with the addition of the larger loci and linkage regions, which were assessed for all of the clinical traits in the family-based association data. The number of loci captured in the family MGA analyses determined the Bonferroni correction factor used. Table 3-4 summarises the multiple testing corrections and p-values used to determine significance for the analysis of the combined 20kb windows and the larger loci

and linkage regions. Significance represents correction for both the number of loci and the number of traits tested for association, whereas suggestive significance represents correction for the number of loci only.

Table 3-4 Significance levels used for analysis of family-based association data for 20kb windows around published SNPs and larger loci and linkage regions.

			p-values	
Regions with reported associations for:	**No. loci represented in association results**	**Association analyses**	**Significance**	**Suggestive significance**
POAG and all linkage regions	101	IOP, VCDR and CCT	1.7×10^{-4}	5.0×10^{-4}
IOP	35	IOP and CCT	7.1×10^{-4}	1.4×10^{-3}
VCDR	30	VCDR	1.7×10^{-3}	
CCT	26	CCT and IOP	9.6×10^{-4}	1.9×10^{-3}
ODA or OCA	29	VCDR	1.7×10^{-3}	

POAG = primary open-angle glaucoma, **IOP** = intraocular pressure, **VCDR** = vertical cup to disc ratio, **CCT** = central corneal thickness, **ODA** = optic disc area, **OCA** = optic cup area

3.3.3 Association of literature reported POAG SNPs with clinical traits in five extended pedigrees

Out of the 41 variants reported in the literature which were captured in the exome sequencing, none were significantly associated with IOP, VCDR or CCT at the Bonferroni corrected significance level of $p \leq 4.1 \times 10^{-4}$ or suggestive significance at $p \leq 0.0012$. Table 3-5 identifies all the variants which reached nominal significance at $p \leq 0.05$ for each of the traits. Five SNPs reached nominal significance for IOP, three for VCDR and one for CCT.

Table 3-5 Nominally significant measured genotype association results for specific SNPs and indels identified in the literature

										Measured genotype association											From Appendix B	
										IOP				VCDR				CCT					
Chr	bp position (hg19)	Variant ID	Gene	Gene region	Ref	Alt	CADD	AF	gnomAD genome NFE	p-value	beta-value	variance explained (%)	mac	p-value	beta-value	variance explained (%)	mac	p-value	beta-value	variance explained (%)	mac	Reference	Associated glaucoma/ endophenotype
2	38,298,139	rs1800440	*CYP1B1*	exonic p.(Asn > Ser)	T	C	**25.00**	0.177	0.170									0.0175	-12.01	7.54	51	[182, 183]	POAG
2	38,302,177	rs1056827	*CYP1B1*	exonic p.(Ala > Ser)	C	A	7.55	0.273	0.306	0.0419	1.48	1.05	135									[182, 183]	POAG
2	38,302,390	rs10012	*CYP1B1*	exonic p.(Arg > Gly)	G	C	5.13	0.273	0.302	0.0419	1.48	1.05	135									[182]	POAG
2	113,542,960	rs1800587	*IL1A*	5' UTR	G	A	2.12	0.321	0.295					0.0354	0.04	2.95	156					[184]	POAG
9	22,003,367	rs1063192	*CDKN2B-AS1*	ncRNA intronic	G	A	0.99	0.655	0.562	0.0435	-1.31	0.85	319									[29, 69, 90, 95-99]	POAG, NTG, VCDR
9	136,131,415	rs8176743	*ABO*	exonic p.(Gly > Ser)	C	T	7.49	0.124	0.105	0.0405	1.89	2.37	61									[86]	IOP
9	136,131,461	rs8176741	*ABO*	exonic	G	A	5.88	0.124	0.105	0.0405	1.89	2.37	61									[136]	IOP, VCDR, ODA or OCA
12	104,415,244	rs11553764	*GLT8D2*	5' UTR	C	T	8.92	0.092	0.173					0.0305	0.08	1.08	45					[179]	CCT
15	100,692,953	rs72755233	*ADAMTS17*	exonic p.(Thr > Ile)	G	A	**23.10**	0.081	0.115					0.0143	-0.09	2.57	38					[171]	IOP

Chr = chromosome, **bp** = base pair, **ref** = reference allele, **alt** = alternate allele
5' UTR - 5' untranslated region, **ncRNA** = non-coding RNA, **exonic** = protein coding region (synonymous unless otherwise indicated)
Ref = reference allele, **Alt** = alternate allele
CADD = combined annotation dependent depletion phred scores (version 1.4), scores ≥ 15 are in bold
AF = frequency of the variant allele in the 5 families, **gnomAD genome NFE** = variant allele frequency in non-Finnish Europeans in the gnomAD genome database
IOP = intraocular pressure, **VCDR** = vertical cup to disc ratio, **CCT** = central corneal thickness
beta-value = effect size per allele. IOP = pressure inside the eye measured in mmHg, VCDR = between -1.0 and 1.0, CCT = thickness of the central cornea measured in μm
mac = minor allele copies as used for MGA analysis (both phenotypic and genotypic data required). Maximum copies possible for each endophenotype are; IOP = 490, VCDR = 486, CCT = 296
POAG = primary open angle glaucoma, **NTG** = normal tension glaucoma, **ODA** = optic disc area, **OCA**, optic cup area
blank cells represent SNPs not reaching nominal significance for the trait

The nine nominally significant variants represent six genes, with *CYP1B1* and *ABO* with more than one associated SNP. All of the variants are common in the general population, as measured by the frequency of the variant allele in the gnomAD genome non-Finnish European database [185], ranging from 10.5% to over 56% in frequency. The effect size of each of the variants was small, as measured by the beta-values. Four of the five SNPs nominally significant for IOP were only associated with average increased pressure of around 1 or 2 mmHg per allele with the *CDKN2B-AS1* variant associated with a reduction in IOP by a similar amount. For VCDR, the effect size was similarly small, with two variants associated with increasing and one variant associated with decreasing VCDR by less than 0.1. The only nominally significant CCT SNP was associated with a decrease in corneal thickness of 12μm per allele. This variant in *CYP1B1* contributes to nearly 8% of the variance in CCT, giving it a larger impact than the other variants, which do not exceed 3% each.

The publications, from which the variants were included in this analysis, are a mixture of smaller targeted gene and variant studies and large GWAS. The *CDKN2B-AS1* variant, rs1063192, is the most replicated variant of these nominally significant associations. Multiple publications associated this variant with POAG, normal tension glaucoma (NTG) and the IOP endophenotype [29, 69, 90, 95-99]. Most of the nominally significant variants in Table 3-5 were identified from POAG loci in the literature. The two *ABO* variants, nominally associated with IOP in the family MGA data were identified from published IOP loci [86, 136]. The two VCDR variants, rs11553764 and rs72755233, were from publications of different clinical traits; CCT and IOP respectively [171, 179].

3.3.4 Association of larger POAG loci with clinical traits in five extended pedigrees

Variants within POAG linkage regions combined with the 20kb windows around published POAG SNPs yielded significant results when tested for association with IOP, VCDR and CCT in the five extended pedigrees. Table 3-6 displays all the variants reaching at least suggestive significance for the three traits. The rarer variants from this table (< 50 copies) are also included in Table 3-7 which supplements allele frequency information with detail on the number of minor alleles present in each of the five families.

Table 3-6 Statistically significant measured genotype association results for 20kb windows around published variants and linkage regions identified in the literature

											Measured genotype association											From Appendix B		
											IOP				VCDR				CCT						
	Chr	bp position (hg19)	Variant ID	Gene	Gene region	Ref	Alt	CADD	AF	gnomAD genome NFE	p-value	beta-value	variance explained (%)	mac	p-value	beta-value	variance explained (%)	mac	p-value	beta-value	variance explained (%)	mac	Region	Reference	Associated glaucoma/ endophenotype
20 kb windows	10	115,805,601	-	*ADRB1*	3' UTR	C	T	10.38	0.010	0.002					3.21×10^{-4}	-0.32	2.18	5					rs1801253 *ADRB1*	[186]	NTG
	11	130,281,493	rs200759153	*ADAMTS8*	exonic p.(Pro > Pro)	G	A	9.11	0.010	0.003	1.16×10^{-3}	7.85	5.59	5									rs4936099 rs55796939 rs56009602 all *ADAMTS8*	[103, 136] [136] [179]	VCDR, ODA/OCA IOP CCT
Linkage regions	2	97,563,745	rs4243786	*FAM178B*	5' UTR	T	C	0.63	0.283	0.264	**1.67×10^{-4}**	-2.12	5.54	138									GLC1B	[175]	POAG
	2	97,877,478	rs10194525	*ANKRD36*	exonic p.(Val > Met)	G	A	**22.50**	0.432	0.419	3.22×10^{-4}	3.44	6.21	211											
	2	98,164,184	rs13001728	*ANKRD36B*	exonic p.(Glu > Asp)	C	G	0.27	0.697	0.679	**7.40×10^{-5}**	2.13	7.49	341											
	2	98,206,188	rs1868445	*ANKRD36B*	5' UTR	C	G	4.71	0.700	0.692	**1.26×10^{-4}**	2.04	7.31	341											
	2	98,206,282	rs28415903	*ANKRD36B*	5' UTR	C	T	5.75	0.699	0.691	**1.25×10^{-4}**	2.00	7.24	342											
	2	98,206,293	rs1814505	*ANKRD36B*	5' UTR	G	A	3.21	0.699	0.691	**1.25×10^{-4}**	2.00	7.24	342											
	2	98,273,290	rs11547232	*ACTR1B*	3' UTR	C	T	0.06	0.685	0.680	**1.06×10^{-4}**	2.06	7.40	335											
	2	197,521,750	rs34133636	*CCDC150*	exonic p.(Glu > Lys)	G	A	**23.20**	0.024	0.062					3.81×10^{-4}	-0.23	6.69	12					2q33.1 – 33.3	[187]	POAG
	3	42,266,360	rs3755809	*TRAK1*	3' UTR	T	C	1.09	0.245	0.215									**5.74×10^{-6}**	24.35	9.30	68	GLC1L	[174]	POAG
	3	49,725,021	rs114429531	*MST1*	exonic p.(Arg > His)	C	T	8.07	0.010	<0.001	**9.60×10^{-5}**	9.49	5.28	5											
	3	126,178,390	rs1043803	*ZXDC*	3' UTR	A	C	4.53	0.608	0.616									4.45×10^{-4}	16.24	10.33	184	GLC1C	[153]	POAG
	3	126,180,121	rs3231	*ZXDC*	3' UTR	C	T	6.93	0.606	0.613									1.96×10^{-4}	16.98	11.36	183			
	3	126,200,403	rs1799398	*UROC1*	3' UTR	A	C	1.41	0.606	0.614									1.96×10^{-4}	16.98	11.36	183			
	5	133,451,564	rs115535840	*TCF7*	5' UTR	G	C	12.85	0.024	0.065									3.60×10^{-4}	-40.62	14.00	11	GLC1M	[48]	JOAG
	5	135,276,202	rs1160982	*FBXL21*	exonic *1 codon deletion*	GTT	-	**19.03**	0.737	0.697									3.48×10^{-4}	-16.90	9.52	218			
	5	149,589,656	rs116239172	*SLC6A7*	3' UTR	C	G	0.64	0.046	0.042									4.25×10^{-4}	-36.48	8.94	10			
	6	167,369,897	rs1044059	*RNASET2*	5' UTR	C	A	10.52	0.398	0.457					**1.30×10^{-4}**	-0.08	8.47	191					6q27	[188]	IOP
	6	167,369,992	rs2247325	*RNASET2*	5' UTR	A	G	9.95	0.281	0.347					**7.90×10^{-5}**	-0.10	8.30	134							
	10	70,742,931	rs12412103	*DDX21*	3' UTR	A	G	6.00	0.064	0.050	**1.20×10^{-4}**	-3.60	6.91	32									10q22	[189]	IOP
	10	72,185,254	rs16927606	*EIF4EBP2*	3' UTR	T	C	4.79	0.036	0.045									3.92×10^{-4}	33.07	5.46	13			

											Measured genotype association												
											IOP				VCDR				CCT				From Appendix B		
	Chr	bp position (hg19)	Variant ID	Gene	Gene region	Ref	Alt	CADD	AF	gnomAD genome NFE	p-value	beta-value	variance explained (%)	mac	p-value	beta-value	variance explained (%)	mac	p-value	beta-value	variance explained (%)	mac	Region	Reference	Associated glaucoma/ endophenotype
Linkage regions	17	6,024,786	rs3809841	*WSCD1*	3' UTR	G	T	0.01	0.413	0.431									**1.20 x 10^{-5}**	-17.73	10.46	128	17p13.3-p13.1	[190]	POAG
	17	8,024,121	rs1442849	*HES7*	3' UTR	C	T	11.90	0.291	0.280									**4.10 x 10^{-5}**	-19.03	11.19	80			
	17	8,701,799	rs17854013	*MFSD6L*	exonic p.(Pro > Thr)	G	T	**16.17**	0.141	0.184	2.43 x 10^{-4}	-2.22	7.04	69											
	17	75,495,397	rs9038	*SEPT9*	3' UTR	T	C	3.99	0.367	0.406	3.20 x 10^{-4}	1.78	3.78	178									17q25.1-q25.3	[190]	POAG
	19	13,010,643	rs9384	*GCDH*	3' UTR	G	T	1.10	0.331	0.360									3.23 x 10^{-4}	-16.01	11.33	102	19p13.2	[191]	IOP
	19	13,044,544	rs2974750	*FARSA*	5' UTR	C	A	6.52	0.659	0.647									**1.11 x 10^{-4}**	17.65	13.11	190			
	20	14,304,700	rs79295487	*FLRT3*	3' UTR	G	A	13.32	0.075	0.012	**1.17 x 10^{-4}**	4.55	7.18	36									GLC1K	[47]	JOAG
	20	14,304,785	rs183080338	*FLRT3*	3' UTR	T	A	**16.07**	0.050	0.008	3.18 x 10^{-4}	5.07	4.73	24											
	20	14,307,024	rs8120693	*FLRT3*	exonic p.(Ala > Thr)	C	T	**19.49**	0.074	0.012	**1.27 x 10^{-4}**	4.54	7.00	36											

Significant p-values indicated in bold, suggestive significance not bold p-values
Chr = chromosome, **bp** = base pair
3' UTR = 3' untranslated region, **5' UTR -** 5' untranslated region, **exonic** = protein coding region, includes amino acid change
Ref = reference allele, **Alt** = alternate allele
CADD = combined annotation dependent depletion phred scores (version 1.4), scores ≥ 15 are in bold
AF = frequency of the variant allele in the 5 families
gnomAD genome NFE = variant allele frequency in non-Finnish Europeans in the gnomAD genome database
IOP = intraocular pressure, **VCDR** = vertical cup to disc ratio, **CCT** = central corneal thickness
beta-value = effect size per allele. IOP = pressure inside the eye measured in mmHg, VCDR = between -1.0 and 1.0, CCT = thickness of the central cornea measured in μm
mac = minor allele copies as used for MGA analysis (both phenotypic and genotypic data required). Maximum copies possible for each endophenotype are; IOP = 490, VCDR = 486, CCT = 296
POAG = primary open angle glaucoma, **JOAG** = juvenile open angle glaucoma, **NTG** = normal tension glaucoma, **ODA** = optic disc area, **OCA**, optic cup area

Table 3-7 Minor allele copies per family for variants identified in Table 3-6 with a < 50 copies genotyped in the five families

										Measured genotype association															
										IOP				VCDR				CCT				Minor allele copy per family					
Chr	bp position (hg19)	Variant ID	Gene	Gene region	Ref	Alt	CADD	AF	gnomAD genome NFE	p-value	beta-value	variance explained (%)	mac	p-value	beta-value	variance explained (%)	mac	p-value	beta-value	variance explained (%)	mac	93001	95002	98002	GTAS04	GTAS54	Total genotyped in 5 families
10	115,805,601	-	*ADRB1*	3' UTR	C	T	10.38	0.010	0.002					3.21×10^{-4}	-0.32	2.18	5					1(1)	1(1)	3(3)			5
11	130,281,493	rs200759153	*ADAMTS8*	exonic p.(Pro > Pro)	G	A	9.11	0.010	0.003	1.16×10^{-3}	7.85	5.59	5									1(1)	1(1)			3(3)	5
2	197,521,750	rs34133636	*CCDC150*	exonic p.(Glu.> Lys)	G	A	**23.20**	0.024	0.062					3.81×10^{-4}	-0.23	6.69	12					4(4)		5(5)	1(1)	2(2)	12
3	49,725,021	rs114429531	*MST1*	exonic p.(Arg > His)	C	T	8.07	0.010	<0.001	**9.60×10^{-5}**	9.49	5.28	5										3(3)	2(2)			5
5	133,451,564	rs115535840	*TCF7*	5' UTR	G	C	12.85	0.024	0.065									3.60×10^{-4}	-40.62	14.00	11	7(6)	3(3)	2(2)			12
5	149,589,656	rs116239172	*SLC6A7*	3' UTR	C	G	0.64	0.046	0.042									4.25×10^{-4}	-36.48	8.94	10	4(3)	2(2)	4(4)	4(0)	9(1)	23
10	70,742,931	rs12412103	*DDX21*	3' UTR	A	G	6.00	0.064	0.050	**1.20×10^{-4}**	-3.60	6.91	32									8(8)		21(21)	1(1)	2(2)	32
10	72,185,254	rs16927606	*EIF4EBP2*	3' UTR	T	C	4.79	0.036	0.045									3.92×10^{-4}	33.07	5.46	13	6(4)	11(8)	1(1)			18
20	14,304,700	rs79295487	*FLRT3*	3' UTR	G	A	13.32	0.075	0.012	**1.17×10^{-4}**	4.55	7.18	36									12(12)	9(9)	13(12)		3(3)	37
20	14,304,785	rs183080338	*FLRT3*	3' UTR	T	A	**16.07**	0.050	0.008	3.18×10^{-4}	5.07	4.73	24										9(9)	13(12)		3(3)	25
20	14,307,024	rs8120693	*FLRT3*	exonic p.(Ala > Thr)	C	T	**19.49**	0.074	0.012	**1.27×10^{-4}**	4.54	7.00	36									12(12)	9(9)	13(12)		3(3)	37

Significant p-values indicated in bold, suggestive significance not bold p-values
Chr = chromosome, **bp** = base pair
3' UTR = 3' untranslated region, **5' UTR -** 5' untranslated region, **exonic** = protein coding region, includes amino acid change
Ref = reference allele, **Alt** = alternate allele
CADD = combined annotation dependent depletion phred scores (version 1.4), scores ≥ 15 are in bold
AF = frequency of the variant allele in the 5 families
gnomAD genome NFE = variant allele frequency in non-Finnish Europeans in the gnomAD genome database
IOP = intraocular pressure, **VCDR** = vertical cup to disc ratio, **CCT** = central corneal thickness
beta-value = effect size per allele. IOP = pressure inside the eye measured in mmHg, VCDR = between -1.0 and 1.0, CCT = thickness of the central cornea measured in μm
mac = minor allele copies as used for MGA analysis (both phenotypic an genotypic data required). Maximum copies possible for each endophenotype are; IOP = 490, VCDR = 486, CCT = 296
POAG = primary open angle glaucoma, **JOAG** = juvenile open angle glaucoma, **NTG** = normal tension glaucoma, **ODA** = optic disc area, **OCA**, optic cup area
Minor allele copy per family = Bracketed minor allele copy per family numbers refer to the number of individuals with phenotype data who contributed to the MGA analysis.

Of the 31 variants identified (Table 3-6), two were from the 20kb windows around the SNP loci from the literature; an unnamed variant upstream from rs1801253 in *ADRB1* on chromosome 10 and rs200759153 which is within the 20kb window of three variants in *ADAMTS8* on chromosome 11. The chromosome 10 variant is in a region identified in a study of NTG [186] and reached suggestive significance in the VCDR association results. This variant is rare in the general population, with a minor allele frequency of only 0.2% and found in a total of 5 heterozygous carriers in 3 of the 5 families (see Table 3-7). This variant appears protective in these families, as it was associated with an average reduction in the VCDR by 0.3. This is quite large, considering that as a ratio, the beta-value for VCDR only ranges between -1 and 0 or 0 and 1, with negative values representing a protective effect and positive values representing a deleterious effect. Although a large effect size, the variance explained in VCDR by this variant was relatively small, at only 2%. The chromosome 11 variant, rs200759153, demonstrated suggestive significance with IOP and is within the 20kb region around *ADAMTS8* variants associated with multiple traits in the literature. Another rare SNP in the general population, at 0.3%, it was also found in 5 heterozygous carriers across 3 families (see Table 3-7). The effect size of this variant was large, with each allele associated with an average increase in IOP of nearly 8mmHg in these variant carriers. The variance of the IOP trait explained by this SNP was also large, at nearly 6%.

The remaining 29 variants of at least suggestive significance are within reported linkage regions. Although larger loci, such as regions of genes, were included in this analysis, associations were not identified within these regions. Most of these variants are within 5' and 3' untranslated regions (UTRs) of genes within the linkage regions. There are 6 non-synonymous protein coding variants and one deletion encompassing a whole codon. Most of the variants are common in the general population, with only 4 variants observed at or below 1% in the gnomAD NFE database.

The first of these rare variants identified within the linkage regions in Table 3-6 is rs114429531, a protein coding SNP in *MST1*, significantly associated with IOP in these families. This variant was extremely rare observed with a frequency of < 0.001 in the gnomAD database. This non-synonymous SNP is carried by 5 family members across two families (see Table 3-7) and increased IOP by an average of 9.5mmHg per allele.

A cluster of rare variants was observed in the *FLRT3* gene on chromosome 20. The three variants were observed at around a 1% frequency in the general population, which increases to

between 5 and 7.5% in the families of this study. Two of the variants are located in the 3' UTR of *FLRT3* and one is a non-synonymous protein coding SNP. These three variants appear to be on the same haplotype, as all the variants are carried by the same individuals in the same families (Table 3-7). One of the three variants, rs183080338, which was the only one of the three to not to reach full significance, is only carried by 25 individuals (see Table 3-7). Twelve members of family 93001 do not carry this variant, but all carry the other 2 variants shared by four out of the five families of this study. This region of *FLRT3* increased IOP between 4.5 and 5mmHg per allele in these families, and explains up to 7% of the variance of this trait. Two out of the three variants have CADD scores greater than 15, predicting deleterious functional effects. The *FLRT3* gene lies in the GLC1K region, originally identified from families with juvenile open angle glaucoma [47].

The remaining significant variants identified in Table 3-6 are more common than those already discussed, ranging from 4% through to 70% in the gnomAD genome NFE database. A group of these common variants are located in the GLC1B region on chromosome 2 and were associated with IOP in this analysis. The first variant in this region, rs4243786, reduced IOP by just over 2mmHg in the families of this study. The remaining variants in this region increased IOP, with rs10194525 demonstrating the greatest IOP increase. Although this non-synonymous protein coding SNP only reached suggestive significance, it has an extremely high CADD score. The other non-synonymous protein coding SNP in this region, rs13001728, is extremely common in both the families and the gnomAD database, at 70% frequency, but has a negligible CADD score. This SNP and the remaining variants in the UTRs of the GLC1B region are found at a similar frequency in the families and have a similar level of effect to IOP, increasing it around 2mmHg on average. Three remaining SNPs reached at least suggestive significance for IOP in Table 3-6; rs12412103, rs17854013 and rs9038. These were found within three different published linkage regions; one on chromosome 10 and two from chromosome 17. Two of these variants are within 3' UTRs and one, rs17854013, is a non-synonymous protein coding variant. This latter SNP reduced IOP in these families, as did the chromosome 10 variant. These two variants are not as common as the third variant, rs9038, which increased IOP by just under 2mmHg in the families.

Three variants reached at least suggestive significance in the VCDR MGA analysis of the linkage regions from the literature. Firstly, rs34133636, a non-synonymous protein coding SNP with a high CADD score in the *CCDC150* gene. This variant was found less frequently in the families than in the gnomAD database and reduced VCDR by 0.23. The remaining two

significant SNPs are within 100bp of each other in the 5' UTR of *RNASET2*. These variants are more common than the other two VCDR variants discussed already and they also reduce VCDR by around 0.1, in these families.

The most significant MGA result in Table 3-6 is for rs3755809, a 3' UTR SNP in the *TRAK1* gene, which was associated with CCT. This is a common variant, with an allele frequency of over 20% in the families and the general population. This variant was associated with an increase in CCT thickness by an average of 24μm and explains over 9% of the variance in this trait in the families. Other significant CCT MGA results were identified in UTRs on chromosome 17. These common variants decreased CCT thickness by around 18μm per allele, in these families. The final significant CCT MGA result is rs2974750 on chromosome 19. This variant increased CCT in these families, whereas another variant within the same published chromosome 19 linkage region, rs9384, decreased CCT at a suggestive level of significance. Both these variants are common in these families. Variants associated with increased CCT at a suggestive significance threshold, were found on chromosome 3 within a 22kb range within the GLC1C linkage region. Variants identified within the GLC1M linkage region decreased CCT measurements between 17 and 41μm, on average, in the families of this study. One of these variants, rs1160982, the only indel in Table 3-6, is a very common deletion in the protein coding region of *FBXL21*. The other two variants in the GLC1M region are much rarer, with less than 5% frequency in the families of this study. They both had a larger effect size of CCT reduction, at around 40μm per allele in the family members, although these results were also at a suggestive level of significance. Another variant with a large effect size is rs16927606 on chromosome 10. This variant was associated with an average increase of CCT by 33μm per allele, but is not as common as other alleles, observed at a frequency of less than 4%.

Table 3-7 shows the rarer variants, from Table 3-6, with less than 50 copies in total across the five families. It should be noted that, as described in section 2.6.3, the minor allele copy number for the MGA analysis was filtered to at least five copies for each variant. Thus, the extremely rare variants have not been included in this analysis. Even the rarest of variants in this analysis, occurring only 5 times across all the families, were found in multiple families.

3.4 Discussion

The major aim of the overall research presented in this thesis was to use extended families enriched for POAG to identify rare variants and propose novel genes which impact the endophenotypes of this complex disease. The bulk of the literature on this disease is focused on common variants with small effect sizes. There have also been large regions of the genome associated with glaucoma, represented by published linkage regions, but with very few functional variants identified from within these. Before identifying potentially novel variants associated with the endophenotypes in the families of this study, it was important to first determine whether variants and regions already recognised in the literature harbour variants influencing POAG risk in these families. There are challenges to comparing the outcomes of published population and family studies with the family generated data from our study. Many different study designs and populations were represented in the literature with the predominant association studies favouring the identification of common variants. Even the family data included in the literature, as published linkage regions, represented multiple family structures and study designs. Most of these published family studies have not found the causative variants or genes. However, as a first stage analysis of our family data, prior to identifying potentially novel POAG variants and genes, using a family-based association test to determine the impact of published variants and regions in our families is a valid approach to take. Thus, recognising that there are limitations to the research conducted in this chapter, MGA analysis was used to assess whether published POAG loci affected trait variance in IOP, VCDR and CCT in the families of this study.

The first stage involved establishing a database of genetic loci (Appendix B) and testing these for association in the five families of this study. Association was analysed for three sets of variant data; published SNPs, 20kb windows around these published SNPs and larger loci, including linkage regions.

3.4.1 SNPs from the literature

Family-based association analysis of the reported SNPs revealed no significant associations (Table 3-5). Fewer than 10% of these variants were present in the WES sequencing in the families, and of these, none reached statistical significance, with only nine reaching nominal significance at $p \leq 0.05$. These were all common variants in the families and in the general population. Most of the variants identified from the family-based association results had small

effect sizes and only accounted for a very small proportion of the variance of the trait. Of the SNPs from the literature reaching nominal significance in this study, one variant, rs1800440 in *CYP1B1*, had a larger effect size and explained a greater amount of variance than the other variants identified. An effect of reducing CCT by an average of 12μm per allele was observed in the families, which is not insubstantial when considering a homozygote for this allele may have a 24μm reduction in corneal thickness. This is a large effect size considering the normal range for CCT is between 503μm and 565μm [192]. This SNP also is ranked in the top 0.1% of deleterious variants by CADD, suggesting it could potentially be of functional significance in these families. *CYP1B1* was originally identified as the causative gene in the primary congenital glaucoma locus GLC3A [50] and since then, has been associated with JOAG and POAG [182, 193-195], although there are conflicting findings on the importance of different variants in the risk of developing POAG [183]. The lack of association of *CYP1B1* with CCT in the literature, including several CCT GWAS [179, 196-201] and the fact that this variant only reached nominally significance in our study, suggest this variant is not an important finding for these families.

When considered in the context of the non-significant p-values, individually the identified variants are unlikely to play an important role in the genetic determination of the POAG traits within these families. However, they may contribute in a polygenic fashion. Although beyond the scope of this study, polygenic risk score (PRS) analysis is becoming a more popular research field and a valuable tool for utilising the rich data obtained from GWAS, which is transferable to the clinical setting [81]. The combined effects of many variants, each with small effect sizes, may provide an increased risk upon which family specific rare variants contribute. PRS studies in large case-control cohorts have recently been published for IOP and POAG [83-85]. Using the risk alleles identified in those studies to determine the background risk of these families would be beneficial to gain a better understanding of the genetic mechanisms important to POAG risk in these families.

3.4.2 20kb around published SNPs

Investigating the regions around the reported variants from the literature may provide other association signals which may be in LD with the variants already identified. Windows 20kb around each variant (10kb upstream and downstream) were chosen to provide an appropriately sized region which would conceivably contain variants in LD with the published SNPs [79]. Only two variants reached suggestive significance within these 20kb window regions; rare

variants in *ADRB1* and *ADAMTS8* (Table 3-6). Although the *ADRB1* variant may be important in the 5 individuals it was identified in, it only explains around 2% of the variance of the VCDR trait, thus is not a major variant in the families overall. This variant was identified downstream from rs1801253, which was implicated in NTG in Japanese glaucoma patients [186]. That study investigated polymorphisms associated with glaucoma and IOP, but not VCDR, which is the associated endophenotype in our study. In the families of our study, the identified variant has a protective effect, associated with an average decrease of 0.3 in VCDR.

The synonymous coding SNP, rs200759153, was identified within the 20kb windows of four *ADAMTS8* variants. These four variants have been associated with each of the clinical traits measured in the literature as well as optic cup area [103, 136, 179]. This SNP was associated with IOP at a suggestive level of significance in this study. This is another rare variant, contributing nearly 6% to the variance in IOP, associated with an average increase in IOP nearly 8mmHg in the families of this study. Several variants have been identified in *ADAMTS8* as associated with the clinical intermediate traits of POAG, and this gene has only very recently been associated with POAG itself [83]. Interestingly, of the five variant carriers of rs200759153 in the families of this study, two had been diagnosed with POAG and two had a suspected POAG diagnosis at the time of examination. The fifth person had not developed POAG at their last clinical examination, but they were only 40 years old at the time and may have developed glaucoma since. As the families of this study are enriched for POAG, this finding may represent an incidental finding. However, this SNP may represent a true association of a variant important in the progression from ocular hypertension to POAG. Further research on the association of this rare *ADAMST8* variant and the gene itself with POAG is needed to elucidate this finding.

3.4.3 Linkage regions from the literature

The published variants themselves and the 20kb windows around them only yielded 2 variants at a suggestive level of significance in the family-based association analysis. The linkage regions identified in the literature were associated with several variants reaching full statistical significance (Table 3-6). These regions total around a tenth of the genome (>300 million base pairs), in comparison to the 20kb windows (< 10 million base pairs) even though a large proportion of these regions had not been sequenced in the WES of these families. The greater number of variants with significant associations in the linkage regions, compared to the 20kb window regions, may be a reflection of the size of the linkage regions themselves, and the

increased chance of observing significant results or an increase in relevance of these regions in the families of this study, or a combination of both of these. Importantly, the significance thresholds determined previously (sections 3.2.3.1 and 3.3.2) were based on a hypothesis of one causal variant per linkage region. We have used that threshold, but in some cases have presented results for more than one significant result per linkage region, for example seven variants from the GLC1B locus (Table 3-6). Although more than one variant may have been identified in a linkage region, they would be considered to be on the same haplotype, with only one causative variant. Haplotype analysis would need to be conducted to confirm these findings, however, in the context of this study, it is hypothesised that only one variant per linkage region per trait would be functional.

3.4.3.1 IOP associated variants within linkage regions

Several significant associations were observed with IOP in these families in four different linkage regions from the literature. GLC1B on chromosome 2 was originally identified in families with POAG with low to moderate IOP levels [175]. The MGA results from our study have identified a haplotype of significant variants for increased IOP, although this increase is only around 2mmHg in the families. Most of these variants are extremely common, at around 70%. Possibly the most interesting variant associated with IOP in the families within the GLC1B region, is rs10194525, a missense mutation with a high CADD score. Although only reaching suggestive significance, this variant had a slightly larger effect size, associated with an average IOP increase of nearly 3.5mmHg per allele, and contributed to over 6% of the trait variance. Of the variants identified within the GLC1B linkage region, rs10194525 is the most likely functional variant in the families of this study.

A variant identified here and worth further consideration is rs114429531 in *MST1*. Extremely rare in the general population, it was associated with a large average increase of 9.5mmHg per allele, in IOP. This variant, observed at 0.2% in the gnomAD NFE database, was identified in five individuals across two families (Table 3-7). Four of the five individuals had a POAG diagnosis, with one person without a POAG diagnosis at their last clinical examination. Similar to the *ADAMTS8* variant already discussed, this correlation of the *MST1* variant with POAG diagnosis may just be a coincidental finding or it may represent a true association of a variant associated with increased IOP and progression to POAG. In a study on differentially expressed genes coding for carbohydrate binding proteins in eyes affected by glaucoma [202], *MST1* was found to be downregulated in POAG trabecular meshwork cells compared to normal cells. As

the trabecular meshwork is critical for aqueous humour outflow, dysregulation of the cells can cause IOP to increase [2, 3, 203]. Further research is necessary to determine whether the *MST1* gene itself is involved with IOP regulation and if so, whether this rare variant affects the functioning of the gene. GLC1L is the linkage region from the literature which encompasses *MST1*. The GLC1L region was identified in a Tasmanian family (different to the Tasmanian families of our study) who also carried a myocilin (*MYOC*) mutation, and association between the linkage region and *MYOC* was proposed [174]. The *MST1* variant and another in *TRAK1* (Table 3-6), discussed with the CCT variants later, are significant in the families of this study, but not likely to be causative variants in the Tasmanian family in which the GLC1L region was identified with an autosomal dominant inheritance pattern. The same family as used to identify the GLC1L linkage region, in a later study was also used to identify a locus on chromosome 10 associated with maximum IOP [189]. In the families of our study, a variant found within this linkage region, rs12412103, was associated with decreased IOP rather than increased. Most of the 32 copies of this variant are carried in Family 98002, one of the American families (Table 3-7).

An association of three variants with within the *FLRT3* gene on chromosome 20 has been identified with IOP. These are uncommon variants (MAF of 0.8% and 1.2% in gnomAD NFE) enriched in the families of this study (MAF of 5% and 7.5% respectively) (Tables 3-6 and 3-7). Two of these variants have CADD scores exceeding 15, with rs8120693, a missense SNP, predicted to be most deleterious of the three variants, with a CADD score of nearly 20. The three variants in *FLRT3* associated with IOP in this study are located in the GLC1K linkage region, one of the identified JOAG loci. JOAG is often more severe with higher IOP levels and greater progression of visual field loss than POAG [203]. GLC1K was originally identified in 25 pedigrees with JOAG [47] and the region refined further with four candidate genes excluded [204]. *FLRT3* was not one of the genes assessed in the Sud et al. (2008) [204] study and the role that this gene may possibly play in IOP regulation is currently unknown.

3.4.3.2 VCDR associated variants within linkage regions

Only three variants reached at least suggestive significance for association with VCDR in the linkage regions from the literature (Table 3-6). A missense variant in *CCDC150* had the highest CADD score (23.2), of all the variants associated within published POAG loci and was identified in 12 individuals across four of the five families of this study (Table 3-7). Although only reaching suggestive significance, it was associated with a reduction in VCDR of over 0.2

and accounted for nearly 7% of the variance of this trait. The other two VCDR associated variants reached full statistical significance and accounted for a larger proportion of the variance (over 8%), but had smaller effect sizes, associated with a reduction in VCDR by an average of 0.1 or less in the family members. These two variants are very common *RNASET2* variants which are within a published IOP linkage region on chromosome 6 [188]. Interestingly, this locus was not replicated in a later study by the same group [191] where they identified another IOP linkage region, this time on chromosome 19. The three VCDR associated variants are unlikely to play an important role in POAG pathogenesis within the families, as the variants are common and associated with a reduction in VCDR, rather than an increase which is observed during POAG pathogenesis.

3.4.3.3 CCT associated variants within linkage regions

CCT associated results were identified across 6 linkage regions from the literature, including POAG, JOAG and IOP loci [48, 153, 174, 189-191]. None of the 11 genes associated with CCT in these families had been implicated with POAG or its clinical traits in those of European descent, as represented by their absence from the POAG loci database (Appendix B). Most of the variants are common, with half of variants associated with a decrease in CCT thickness in these families. As thinner corneas are associated with increased risk of POAG [2, 13, 178], variants associated with a thinner CCT are more relevant to POAG pathogenesis. Although only reaching suggestive significance, two less common variants; rs115535840 and rs116239172, within the GLC1M linkage region had large effect sizes and accounted for a large proportion of the trait variance (Table 3-6). The first variant was associated with an average 41μm thinner cornea and accounted for 14% of the CCT trait variance in these families and the second; with an average 36μm thinner cornea and accounting for 9% of the variance. Unfortunately, although the latter variant in the *SLC6A7* gene was identified in 23 individuals across all of the families, CCT phenotyping data was only obtained for 10 of these family members, meaning association testing could only be undertaken on these 10 individuals (Table 3-7). It would be interesting to determine if having the full CCT phenotyping data would strengthen the observed association signal. The GLC1M locus, within which these variants are located, was identified in a single family from the Philippines with many family members affected by JOAG [48]. This locus was refined further and *NRG2* excluded as a candidate gene [205]. It is unlikely that the variants associated in our study were causative variants in the

original study, but they are still of interest and may be involved with determining CCT in the families here.

The four statistically significant CCT associations in Table 3-6 were identified within three linkage regions from the literature and were associated with both thicker (rs3755809 and rs2974750) and thinner (rs3809841 and rs1442849) corneas. They are all common variants in the general population with moderate effect sizes of ±18μm to 24μm and accounted for 9% to 13% of the CCT variance. These variants, as well as the ones only reaching suggestive significance, may be true genetic determinants for CCT, but may not be involved in POAG pathogenesis in the families of this study, as CCT is not genetically correlated with POAG in these families [42]. It is still important to identify genetic determinants of CCT to gain a better understanding of the biological pathways important in corneal development, as thinner corneas are associated with a greater risk of developing glaucoma [2, 13, 178], other corneal diseases such as keratoconus [179] and progression of ocular hypertension to POAG [206, 207].

3.4.4 Contribution to trait variance

CCT is the most heritable of the three traits (Table 3-3), and as such there is more variance available, allowing for larger effects of the identified variants. The CCT variants identified in this study demonstrated this, with the greatest beta-values and with the most contribution to the trait variance when compared to the IOP and VCDR variants (Table 3-6). Across all of the traits in this study, the variants with the largest effect sizes were the rarer ones, occurring less than 1% in the general population. Chen et al., 2014 also found stronger family-based association results with rarer variants [208]. Considering how rare some of the variants are, the proportion of variance to which they contribute is large. For example, the IOP associated variant in *MST1* (rs114429531) was only identified in five individuals but contributed to over 5% of the IOP trait variance, which is a similar proportion to the contribution of rs4243786 in the *FAM178B* gene, which has 138 copies of the allele in the families. The largest contribution to trait variance of a single variant is rs115535840 in *TCF7*, in which 11 copies of this allele contributed to 14% of variance in CCT.

3.4.5 Strengths and limitations

A limitation of this study is CCT measures were only available for a subset of family members (148 sequenced individuals; primarily from the American families), due to these measurements

only being implemented as a routine clinical measurement several years after the initiation of these studies. This decreased the power of the family-based association analyses for this trait, although this was somewhat balanced by the boost from the increased heritability. However, this was for the CCT trait only. Most of the family members with genotypic data also had complete IOP and VCDR phenotypic data (see Table 2.1 Figures 2-1 to 2-5). The MGA analysis itself was run as an additive model, with the heterozygous genotype considered halfway between the two homozygotes [167]. This may be an appropriate model for many variants, but is unlikely to be the best model for every variant tested. As the association analysis was performed on each variant independently, without any form of conditional analysis, interpretation of the results needs to be considered carefully. Variances cannot be added together or assumed to be independent for different loci, as there may be interactions between loci. For example, the cluster of IOP associations on chromosome 2, within the *ANKRD36B* and *ACTR1B* genes (Table 3-6), are assumed to be a haplotype, with an average pressure increase of 2mmHg and around 7% contribution to the trait variance for the whole haplotype, not each individual variant. However, without testing, it cannot be assumed that this haplotype is independent of other IOP associations identified in the table.

The purpose of the study presented in this chapter was to determine whether POAG associated loci from the literature contribute to trait variance in the families of this study. An important limitation of this study is that only the protein coding regions as captured by exome sequencing could be assessed. To overcome this, whole genome sequencing would be required as the majority of associations detected by GWAS are in non-coding regions [79]. SNP array analysis could also be conducted, followed by imputation to provide similar data to that analysed in GWAS [209, 210]. However, increased genotyping of variants would be accompanied by an increase in the multiple testing penalty. The fact that only 10% of the SNP loci from the literature were represented in the family genotype data demonstrates the inability to determine the contribution of non-coding variants to trait variance in these families. However, the financial considerations limiting sequencing to exomes only, allowed for many more family members to be sequenced than if whole genome sequencing was conducted instead.

A strength of this study is the size and complexity of the families and quantity of family members who have contributed both phenotypic and genotypic data. The use of families allows for the identification of rare variants due to enrichment of these variants as they segregate through the generations of the family [211]. Although some of the rare variants presented in this chapter have demonstrated enrichment in the five families overall (Table 3-7), they are not

confined to a single family, but spread over multiple families. For example, the *MST1* variant, which is extremely rare in the gnomAD genome NFE database (MAF < 0.001), was present in two of the five families. Due to the rarity of this variant, this result is unexpected. Although this may be a genuine result in these families, it may actually represent a sequencing error and all of the variants would need validating by sequencing in the laboratory before being prioritised further. Some of the less rare variants, for example the three FLRT3 variants, which are also enriched in these families when compared to the gnomAD database (Table 3-7), provide confidence that genuine enrichment has occurred as it would be unlikely to obtain erroneous sequencing results across 3 variants.

3.4.6 Conclusion

Overall, the study presented in this chapter did identify variants affecting trait variance for IOP, VCDR and CCT. However, individually, some of these effects were small and unlikely to be the main contributors to trait variation and POAG risk in these families. A future PRS analysis may determine the underlying risk these common variants contribute in the families of this study. There were rare variants identified with larger effect sizes and these need to be validated in the laboratory before prioritising further. Variants worth further consideration are rs114429531 in *MST1* and the *FLRT3* region for IOP, and rs115535840 (*TCF7*) and rs116239172 (*SLC6A7*) for CCT. Recognising that IOP is a true POAG endophenotype but CCT is only a risk factor means that the CCT variants identified may not be involved with POAG pathogenesis, but may still be useful to gain an understanding of corneal development. As IOP is currently the only modifiable risk factor for POAG and the primary target for treatment, it is important to identify genes which may be implicated with this POAG endophenotype. A better understanding of the pathways leading to increased IOP will ultimately lead to better treatments.

Chapter 4
Variance components linkage analysis of whole-exome sequencing data in five extended pedigrees

---

4.1 Introduction

Conducting linkage analysis using quantitative intermediate traits of POAG has the potential to identify genes important in POAG pathogenesis [37, 42, 68, 117]. Quantitative traits provide objective data for analysis, which is more powerful than a more subjective dichotomous disease status based on reaching threshold trait values. Early linkage studies focussed on this discrete diagnosis of glaucoma disease status [46, 153, 175] but more recently, the value of using quantitative traits has been recognised [42, 133, 188, 191, 212]. Phenotypic variation in the quantitative trait due to genotypic variation is necessary to identify a quantitative trait locus (QTL), the region of DNA where the genotypic variation occurs. The quantitative traits in this study have previously been assessed as suitable endophenotypes for glaucoma [42], with IOP and VCDR demonstrating statistically significant correlation with POAG. This correlation with POAG is important, as the purpose of this research is to identify QTLs which are not only important in the trait itself, but also involved in POAG pathogenesis.

Linkage analysis which uses variance components methods are suitable for quantitative traits involved with complex disease, such as glaucoma [162]. Multiple loci can be assessed in large, complex pedigrees, with larger pedigrees providing more power for linkage than smaller pedigrees [167, 213]. Variance components methods partition the phenotypic variation of a trait into its components consisting of an additive polygenic component and random environmental effects component, as shown in equation (1).

$$\sigma^2{}_{\mathrm{p}} = \sigma^2{}_{\mathrm{g}} + \sigma^2{}_{\mathrm{e}} \quad (1)$$

Variance is represented as σ^2 with p, g and e denoting phenotypic, genetic and environmental components respectively. Equation (2) is an extension of the basic variance components model with covariance of family members included. Covariance takes into account the specific relationships between family members.

$$\Omega = 2\phi\sigma^2{}_a + I\sigma^2{}_e \quad (2)$$

The phenotypic covariance, Ω, is partitioned into its additive genetic and environmental components. The genetic component represents the contribution of additive genetic variance anywhere in the genome ($\sigma^2{}_a$) including the coefficient of relationship (2ϕ), describing the familial relationship between each pair of individuals. Phenotypes are expected to be more similar the more closely related the individuals. The environmental component, which is expected to be unshared in large families, is reflected with the included identity matrix (I).

To conduct linkage analysis within a variance components methodology, a further extension to equation (2) is needed. Equation (3) demonstrates this additional level of allele sharing information.

$$\Omega = \Pi\sigma^2{}_{qtl} + 2\phi\sigma^2{}_a + I\sigma^2{}_e \quad (3)$$

Rather than only including an additive genetic effect, which encompasses any number of loci across the genome, QTL specific variance is incorporated ($\sigma^2{}_{qtl}$) which accounts for allele sharing information between relatives. Identity by descent (IBD) matrices describe the measured amount of allele sharing between individuals in a family; the closer the relationship, the more allele sharing at a locus is expected. IBD measures shared alleles which are inherited, without mutation or recombination, from a common ancestor. This is in contrast to measuring identity by state (IBS), where relatives might share the same alleles, but these alleles may have entered the family from multiple founders. Alleles inherited IBD are expected to be rarer than those inherited IBS, as they must be inherited from a common ancestor without any alteration. Alleles inherited IBS can potentially come from multiple family members and are not as useful for localising QTLs as the information gained from alleles inherited IBD. The proportion of alleles inherited IBD at the specific QTL is represented by the Π matrix for that genomic location. Residual additive genetic effects, outside of the QTL, are represented by $2\phi\sigma^2{}_a$ in equation (3). In this way, QTL specific variance is partitioned from the total genetic variance and allows for the identification of which QTLs contribute most to the total trait variance.

An accurate measure of IBD is crucial for successful linkage analysis with variants components methods [162]. Multipoint IBD estimation, where IBD at a location is calculated based on surrounding marker information, is better than single point estimation for identifying QTLs using variants components linkage analysis [214]. Early linkage studies used around 400

microsatellite markers across the whole genome for IBD estimation. In contrast, SNP arrays and next generation sequencing technologies such as WES can produce many thousands of markers. Microsatellite markers were originally chosen because of their high heterogeneity due to varying allelic lengths, in contrast to biallelic SNP markers which are a single base pair in length, restricting their allelic diversity. The information content provided by these markers is critical to successful IBD estimation and linkage analysis [215]. Research comparing the use of microsatellite markers with SNP array markers has determined that a higher density of SNP markers is comparable to and may exceed the use of lower density microsatellites in linkage analysis [216-218]. However, the LD of these high density SNPs needs to be considered, as SNPs in high LD may inflate LOD scores [218].

Conducting linkage analysis from WES data, rather than SNP array data, provides the advantage of potentially being able to identify any functional exonic variants themselves. Gazal et al. (2016), compared the performance of linkage analysis conducted using WES sequencing with that using SNP arrays and found them similar, with the WES sequencing proving to be more cost effective in being able to reduce the numbers of candidate variants than using SNP arrays [219].

Different methods and algorithms need to be used when calculating IBD from high density SNP data than microsatellite data. Computationally, algorithms which are suitable for IBD estimation with few markers or small families may not be suitable for the dense data generated from WES in large, extended families [220, 221].

4.1.1 Aim of this study

To identify quantitative trait loci for intermediate POAG traits using linkage analysis in five extended pedigrees.

Specifically, identity by descent will be estimated from whole exome sequencing data and used to conduct variance components linkage analysis on five POAG enriched families to identify quantitative trait loci for IOP and VCDR.

4.2 Methods

4.2.1 Preparation of WES data for multipoint linkage analysis

Sequencing and variant calling of the exome data was discussed in Chapter 2. The variant call format (VCF) file generated contained 435,975 biallelic SNPs and was the starting point for preparation of files suitable for conducting variance components linkage analysis. Only biallelic SNPs were used for identity by descent estimation, due to the requirements of the chosen program. The steps taken to prepare the VCF data for linkage analysis are summarised in Figure 4-1. Each of these steps required optimisation and are detailed in the following sections.

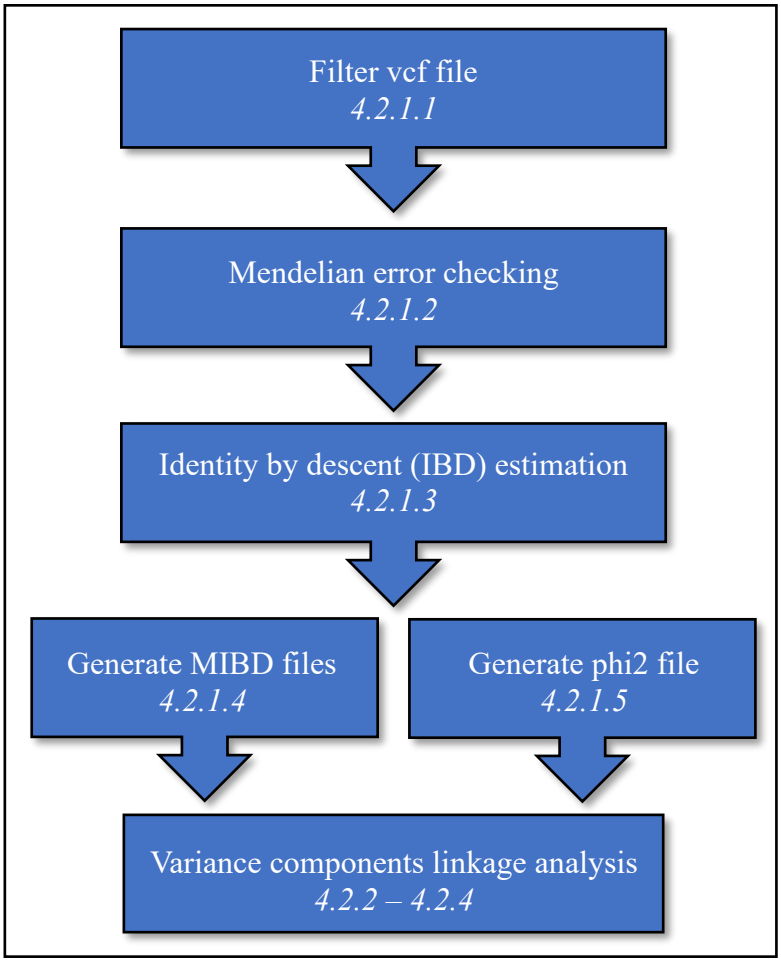


Figure 4-1 The major data manipulation steps prior to variance components linkage analysis of WES data. Italicised numbers refer to the following sections with detailed descriptions of each step.

4.2.1.1 Filter VCF file

Biallelic SNPs were extracted from the full VCF file obtained after variant calling. This VCF file was first labelled with quality annotations using the GATK "Variant Filtration" tool (as described in Chapter 2.4.2). Only variants with "high-confidence" calls were selected for linkage analysis. A custom R script was used to set the poor quality variant calls to missing (see Appendix D.1). This same R script was used to calculate the number of observations of each variant. Variants were deleted if observed less than eight times across all five families. Inclusion of very rare SNPs may affect identity by descent (IBD) estimation and subsequent linkage analysis [222]. Nearly three-quarters of the SNPs were filtered out in this stage. This figure includes variants which were genuinely rare and also variants which were artificially rare due to the blanking out of variants annotated as "low".

Only the autosomal variants were considered in our study and the different methods which are required for the sex chromosomes is for a future study. PLINK (version 1.9) [223] was used to remove X and Y chromosome variants. PLINK was also used to remove variants with missing genotypes exceeding 5%. This threshold was chosen as it was the same threshold used for identity by descent estimation (see section 4.2.1.3). At the same time, variants which deviated considerably from Hardy-Weinberg equilibrium were also removed. Only founders were considered in this test. A threshold of $p < 0.001$ for the Hardy-Weinberg equilibrium exact test was used, as it allowed for a slight deviation from Hardy-Weinberg equilibrium, which may occur with genuine associations of variants with traits. Appendix D2 shows the PLINK commands used for this stage of the filtering.

Table 4-1 summarises the VCF file filtering stages discussed above. At the conclusion of filtering the initial VCF file, 67,605 biallelic SNPs remained, all of high-coverage (depth ≥ 10) and high-quality (genotype quality score ≥ 20).

Table 4-1 Filtering stages of the initial VCF file

Filtering process	Filtering threshold	Number of SNPs removed	Number of SNPs remaining	Filtering program
Initial VCF file			435,975	
Remove very rare SNPs	MAC < 8	321,823	114,152	R script to set poor quality calls to missing and to filter out rare SNPs
Remove X and Y chromosome SNPs		2,347	111,805	PLINK
Remove missing genotypes	> 5% missing	42,153	69,652	PLINK
Remove SNPs which deviate from Hardy-Weinberg equilibrium	p-value < 0.001	2,047	67,605	PLINK

MAC = minor allele copy

4.2.1.2 Mendelian error checking

The filtered VCF file was checked for Mendelian inconsistencies using the PedCheck program (version 1.1) [224]. PedCheck was run for each chromosome individually, with loops to automate the process (see Appendix D.3). The process of file preparation for PedCheck is detailed in Figure 4-2.

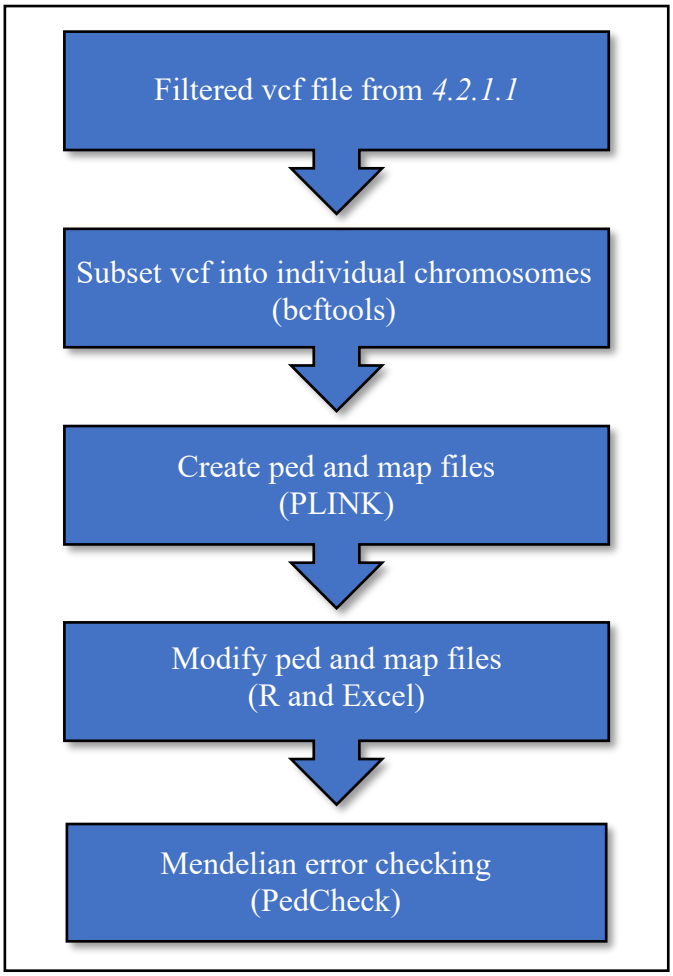


Figure 4-2 Preparation of files for Mendelian error checking. Programs used for each stage are in brackets.

Firstly, bcftools (version 1.8) [225] was used to subset the filtered VCF file into VCF files of individual chromosomes (see Appendix D.3.1). PLINK was used to create ped and map files from these individual chromosome VCF files (see Appendix D.3.2) The PLINK produced map files were concatenated, adjusted to 1 column with a unique chromosome_base-pair identification number and then subset again into individual chromosome map files (see Appendix D.3.3). The PLINK produced ped files were adjusted to the requirements of the

PedCheck program using R. As there was one ped file per chromosome, this step was streamlined using a loop within the R script (see Appendix D.3.4).

After file preparation, PedCheck was used to run Mendelian error checking (see Appendix D.3.5). PedCheck was used to check for inconsistencies between parent-offspring relationships, as well as inconsistencies beyond the basic nuclear family, by using a "genotype-elimination" algorithm recursively [224]. Results of Mendelian error checking, with the number of markers with errors per chromosome, are presented in Table 4-2.

Table 4-2 Mendelian errors identified per chromosome

Chr	Total no. SNPs	No. markers with errors	% markers with errors
1	6960	187	2.7
2	4494	79	1.8
3	3645	63	1.7
4	2872	36	1.3
5	3324	46	1.4
6	3201	16	0.5
7	3613	73	2.0
8	2610	52	2.0
9	2714	54	2.0
10	3015	39	1.3
11	4196	87	2.1
12	3501	51	1.5
13	1363	9	0.7
14	2162	48	2.2
15	2255	55	2.4
16	2963	35	1.2
17	3878	55	1.4
18	1203	18	1.5
19	4767	117	2.5
20	1978	22	1.1
21	987	11	1.1
22	1904	36	1.9
Total	**67605**	**1189**	

Variants identified as Mendelian errors were excluded from all five families, even if the error did not occur in all of the families, to allow for consistency of variants used for linkage analysis. Names of the error SNPs were extracted from the PedCheck output and PLINK was used to remove these SNPs from the previously filtered VCF files (see Appendix D.3.6).

After removing the identified variants from the VCF files, the Mendelian error checking procedure was repeated, to confirm that no Mendelian inconsistencies still existed. Once this was confirmed, the individual chromosome VCF files were concatenated into one full VCF file, using bcftools. After Mendelian error checking, 66,416 variants remained.

4.2.1.3 Identity by descent (IBD) estimations

IBD estimations were performed using IBDLD (version 3.38) [152, 226]. This program is designed for dense genotype data and estimates IBD allele sharing between each pair of individuals in all of the pedigrees. An extension of the hidden Markov model is used for multipoint IBD estimation which includes calculations of linkage disequilibrium (LD), thus LD pruning of dense genotype data is not necessary. The GIBDLD method, within IBDLD, was recommended over the similar LD-RR method, due to a known issue with distant relationships in the latter method (Mark Abney, personal communication). The GIBDLD method does not require detailed pedigree information and is able to calculate IBDs between members within the same pedigrees and between pedigrees.

Preparation of the files for IBD estimation is described in Appendix D.4. Firstly, PLINK was used to create map and ped files from the corrected filtered VCF file (see Appendix D.4.1). PLINK used the SHAPEIT recombination map (https://mathgen.stats.ox.ac.uk/genetics_software/shapeit/shapeit.html#gmap) to include centi-Morgan (cM) positions for each of the variants in the map file output. The map file was then adjusted using Excel, to meet the input requirements of the IBDLD program (see Appendix D.4.2). A custom R script (Appendix D.4.3) was used to modify the PLINK output ped file to the appropriate pedigree file format for IBDLD.

The IBDLD program required optimisation for several parameters and the full command is provided in Appendix D.4.4. Optimisation of the parameters was based on the detailed information provided in the IBDLD software documentation with additional recommendations provided by the authors (personal communication). As only variants of high coverage and high confidence were used (see section 4.2.1.1), a minimum call rate of 95% per variant was set as the threshold for IBD computation. The `-rars` ("remove adjacent redundant SNPs") parameter was used to ensure that the remaining SNPs all contributed to IBD estimation. This parameter removed SNPs in full LD with each other which did not provide additional information for LD modelling. A total of 241 SNPs were excluded in this initial stage, leaving 66,175 SNPs remaining for IBD estimation. The `-ploci` and `-dist` parameters were used for the LD

calculations prior to IBD estimations. LD was calculated from one end of the chromosome to the other (smallest base pair position through to largest), and `-ploci 50 -dist 2` used 50 previous SNPs within a 2cM window for these LD calculations. The `-hiddenstates` parameter was set to 9, which is the appropriate setting for relatedness estimations in large pedigrees. The `-ibd` parameter was set to 2, which outputs IBD sharing at each locus as well as the Δ7 of the nine condensed identity coefficients, as defined by Jacquard (1972) [227]. The Δ7 coefficient refers to the probability of both alleles being inherited IBD at each locus and is a requirement for the use of the multipoint IBD files in the SOLAR program. IBDLD was run on the 66,175 filtered and Mendelian error corrected SNPs. IBD estimation was conditional on all of the SNPs on the chromosome and output files for each chromosome were obtained. IBD estimates for each chromosome were output as zipped text files and these were used to generate the multipoint identity by descent (MIBD) matrices.

4.2.1.4 Generate multipoint identity by descent (MIBD) matrices

Output from the IBDLD program was converted to MIBD matrices specific for the linkage analysis protocol. A custom awk script (see Appendix D.5.1) was used to extract the multipoint IBD data from the zipped IBDLD output files which were generated per chromosome. The map file required for this process was a modification of the map file used for the IBDLD program. The map file for IBDLD estimation was too dense to use for linkage analysis and the density of markers was reduced to 1 SNP per cM (see Appendices D.5.2 and D.5.3 for the modification details).

The awk script used for generating MIBD files (Appendix D.5.1) also required access to a numerical index of the study participants, consistent with the indexing which occurs when linkage analysis is conducted. These index files were generated when the pedigree file was loaded into the linkage program. More details about the pedigree file and linkage program are provided in section 4.2.2.

To generate MIBD files, the custom awk script (Appendix D.5.1) was used to combine IBD information from IBDLD output files with pedigree index numbers and marker locations. A loop command was used to streamline this process (Appendix D.5.4). One MIBD matrix was generated for each marker in the map file. A further processing step of the MIBD matrices was conducted, which scaled the MIBD matrices to ensure individual values did not greatly exceed 1. A custom C++ program, grm (version 0.3-48-g7550b77, written and provided by Juan Peralta, University of Texas Rio Grande Valley), was used for this purpose using the

approaches previously described [228, 229]. The command to run this program is detailed in Appendix D.5.5. The default tolerance threshold of $1x10^{-6}$ was used to ensure that no MIBD value exceeded 1.00000. Use of the MIBD matrices is detailed further in section 4.2.2.

4.2.1.5 Generate an empirical kinship (phi2) file

An empirical kinship matrix file, containing accurate family relationships, was necessary for the linkage analysis. As the IBDLD program was used to determine pairwise IBD, from which MIBD matrices were derived, this program was also used for the empirical kinship calculations. As well as producing IBD estimations per chromosome, the IBDLD program also generated genome-wide empirical kinship data. Appendix D.6 outlines the preparation of the phi2 file. Details of the use of the phi2 file are provided section 4.2.2.

4.2.2 Multipoint linkage analysis with ascertainment correction

Variance components multipoint linkage analysis was conducted using SOLAR [162] with the MIBD and phi2 matrices described above (section 4.2.1.4 and 4.2.1.5 respectively). Linkage was conducted on the IOP, IOPmed and VCDR traits (see section 2.5.1 for description of traits) on 249 members of the five families. CCT was not included in this analysis as there were not enough individuals with CCT measurements in each of the families to conduct a robust analysis.

The pedigree file used did not provide parental or family information (these 3 fields were zeroed out) for each of the 249 sequenced individuals of this study. Thus only each individual's unique identification number and sex were included in the pedigree file. This allowed for the use of the phi2 file to provide familial information based on calculated empirical kinship values from the IBDLD program (section 4.2.1.5). The phenotype file also provided personal and trait data on only the 249 sequenced individuals. A null model, a model with no linkage elements included, was first created with polygenic modelling. The modelling for linkage analysis differed from the polygenic modelling described in Chapter 2 prior to MGA (Chapter 2.6.1) as that modelling used the full pedigree file, including all family relationships, from which SOLAR generated a kinship matrix to conduct MGA. The linkage analysis discussed in this chapter used a pedigree file without family relationships with the IBDLD-generated empirical kinship matrix.

The linkage analysis was conducted with ascertainment correction. An ascertainment correction is used to correct for a potential distortion of data when families are non-randomly ascertained. As discussed previously (Chapter 2.2) the five families from OGGS and GIST recruited in this study were enriched for POAG. Their phenotypic measurements, such as IOP, may not represent measurements as observed in the general population. Mean IOP measurements may be higher in the families than the general population and the variance may be lower, if there is less representation of individuals with normal IOP in the families. By constraining the linkage analysis to the population means, including effects of covariates, a more accurate representation of the linkage patterns of that trait may be observed.

To provide a population ascertainment correction for linkage analysis conducted on the POAG endophenotypes, data from the Melbourne Visual Impairment Project (MVIP) was used [230]. This population based study provided data for participants aged 40 and over and clinical examinations including IOP and VCDR measurements were conducted in a similar manner to those of the OGGS and GIST. Table 4-3 compares age, IOP and VCDR data between the families involved with this study and the MVIP population. Overall, the families of this study were younger than the MVIP participants, had higher IOP and slightly larger VCDR.

Table 4-3 Comparison of mean age, IOP and VCDR measures between the families of this study and the Melbourne Visual Impairment Project

	Families	MVIP
Mean age (years)	54.68 ± 19.21 (312)	58.19 ± 12.35 (4363)
Mean IOP (mmHg)	20.18 ± 2.51 (301)	15.43 ± 3.34 (3691)
Mean VCDR	0.44 ± 0.25 (301)	0.45 ± 0.20 (3451)

IOP = intraocular pressure, **VCDR** = vertical cup to disc ratio,
each mean is ± the standard deviation
Numbers in brackets refer to the number of participants with data available

Polygenic analysis of the MVIP data provided the beta-values (coefficients) for the covariates, as observed in an unrelated, generally healthy population. Data from this population was used to constrain the mean for each trait, as well as the interactions of the covariates. The values to be constrained were included in a tcl (tool command language) script within SOLAR (see Appendices D.7 and D.8). The polygenic modelling was conducted with these constraints in place and the resultant null model used for the linkage analysis. By constraining these values

for multipoint linkage analysis, a more realistic indication of where the family data sits, in relation to the general population, is obtained.

To conduct the linkage analysis, the trait was specified and chromosomes selected (`chromosome all` or `chromosome 1-22` will run linkage across all autosomes). The interval between markers can also be selected, with an interval of 1 meaning that each marker in the map file will be used. For this analysis, an interval of 1 was used to gain maximal linkage information (at a maximum of 1 SNP per cM). The multipoint scanning was first set at `multipoint 3 2`, meaning that a first pass across the genome searched for the largest QTL with a minimum LOD of 3 or more. If a QTL meeting this criterion was identified, a second pass across the genome was undertaken, with largest QTL fixed (removed) from the subsequent pass. The QTL specific variance, is represented by $\Pi\sigma^2_{qtl}$ in equation 3 in the introduction to this chapter. It is partitioned from the total genetic variance and allows for QTLs to be identified individually in successive passes. This continued until the LOD of the largest QTL did not reach 2. The commands used to conduct the linkage analysis using SOLAR are provided in Appendix D.9.

4.2.3 Empirical LOD adjustment

Variance components linkage analysis is sensitive to the distribution of the data, where a non-normal distribution may amplify the size of a linkage peak by increasing type 1 errors [212, 231]. An empirical LOD adjustment can be performed on non-normally distributed data and provides a constant, which when multiplied by the observed LOD scores, gives a more accurate indication of the expected LOD score under a normal distribution. A constant of less than 1 indicates the data is inflated and requires adjusting. After polygenic modelling and ascertainment correction, a LOD adjustment calculation was run with 10,000 replications for each trait with high kurtosis. The `lodadj -calc` command within SOLAR simulates a distribution of LOD scores which would be expected under the null hypothesis of no linkage. Each replication calculates a LOD for a simulated marker unlinked to the trait. SOLAR conducts a regression analysis of the observed LOD scores against the simulated LOD scores providing the LOD adjustment constant. If this constant is less than 1, it is multiplied by the observed LOD scores to more accurately reflect the LOD scores which would be obtained had the distribution been normal.

As the lodadj function in SOLAR does not currently support empirical LOD adjustment using the ped file without family associations (as described in section 4.2.4), the full ped file, with the all family and parental information, was used for this analysis. Empirical LOD adjustment was conducted on the IOP and IOPmed traits.

4.2.4 Significance of linkage analysis

Linkage plots were drawn highlighting both significance and suggestive significance levels. A LOD score of 3.3 was used for full significance and 1.86 used for suggestive significance, as proposed by Lander and Kruglyak (1995) [232].

4.3 Results

4.3.1 Polygenic modelling of zeroed ped file

Polygenic modelling for each of the traits used for linkage analysis are shown in Table 4-3. The same ped file which was used to generate the IBDLD matrices (section 4.2.1.3) was used for this modelling. All of the traits were significantly heritable, with IOPmed more heritable than the unadjusted IOP trait by an order of magnitude in the p-value. Age was a significant covariate for each of the traits, with age^2 also significant for both of the IOP traits. Both of the IOP traits were kurtotic, indicating a non-normal distribution of phenotypic measures.

Table 4-4 Variance components polygenic modelling of POAG clinical traits used for linkage analysis.

Trait	h^2r ± SE (%)	p-value	Residual kurtosis	Covariates included in final model (non-normalised traits only)	Proportion of variance due to covariates (%)
IOP	43 ± 11	1.20×10^{-6}	**3.44**	age, age^2	20
IOPmed	50 ± 11	2.00×10^{-7}	**2.32**	age, age^2	27
VCDR	41 ± 13	3.91×10^{-5}	-0.55	age	33

POAG = primary open-angle glaucoma, **IOP** = intraocular pressure, **VCDR** = vertical cup to disc ratio, **CCT** = central corneal thickness, **norm** = normalised trait,
h^2r = heritability, **SE** = standard error, **residual kurtosis** = a measure of non-normal distribution, bold indicates kurtosis exceeding 0.8 threshold

4.3.2 LOD adjustment

LOD adjustment was conducted for the ascertainment corrected IOP and IOPmed traits, due to their high kurtosis (see Table 4-3). The resultant adjustment constants were 0.63 for IOP and 0.59 for IOPmed. LOD scores from the linkage analyses were multiplied by these constants to adjust for inflation due to the non-normal distributions of these traits. LOD adjustment was not necessary for the VCDR trait.

4.3.3 Multipoint linkage analysis for intraocular pressure

Variance components multipoint linkage analyses for the ascertainment corrected IOP and IOPmed traits are shown in Figure 4-3, including the application of LOD adjustments which reduced the linkage peaks accordingly.

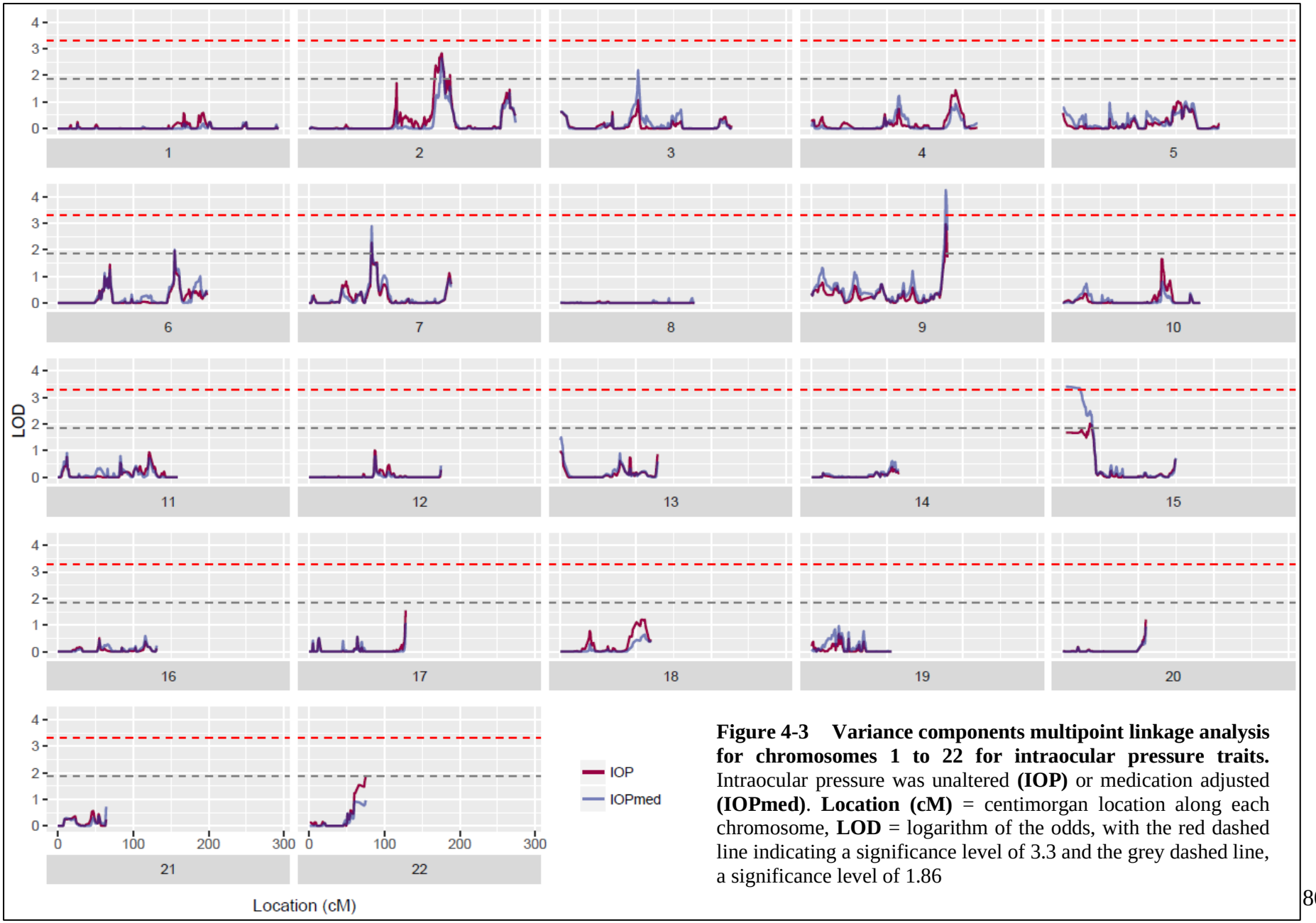
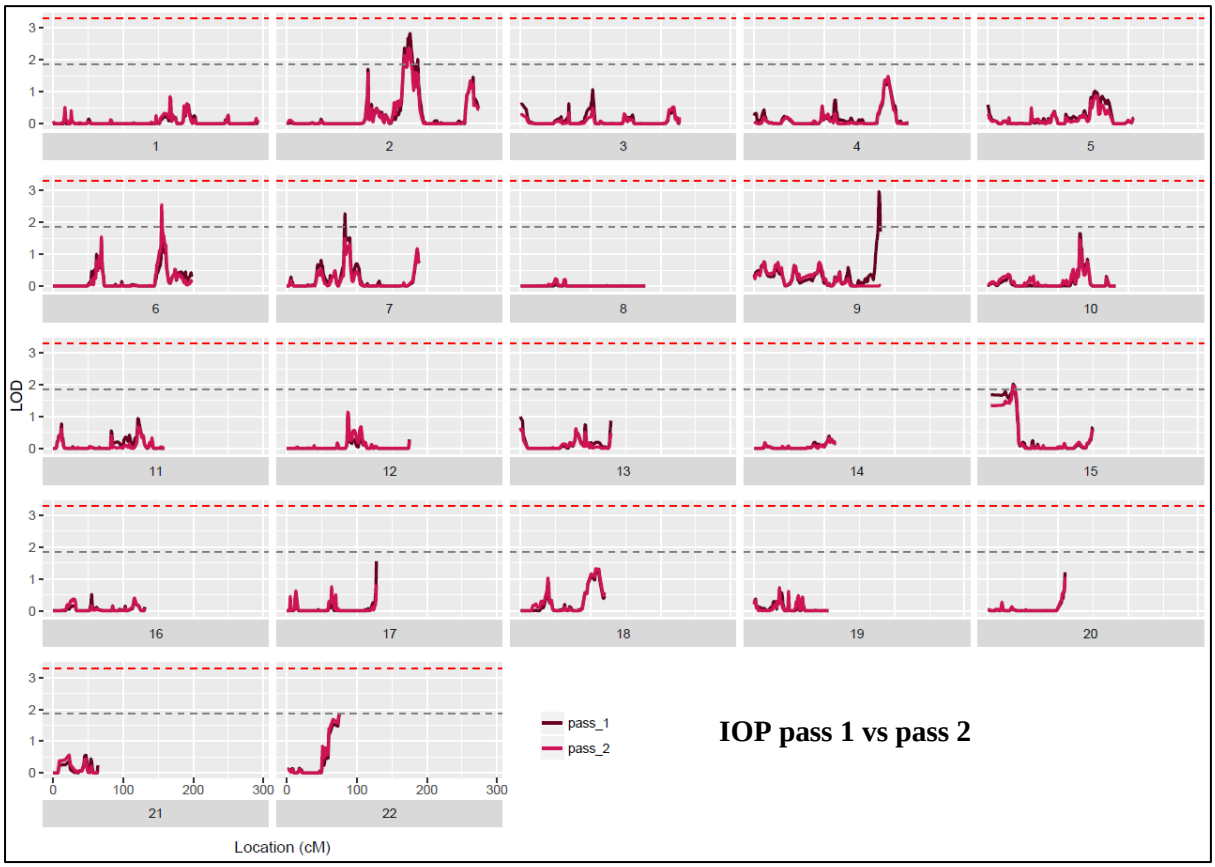


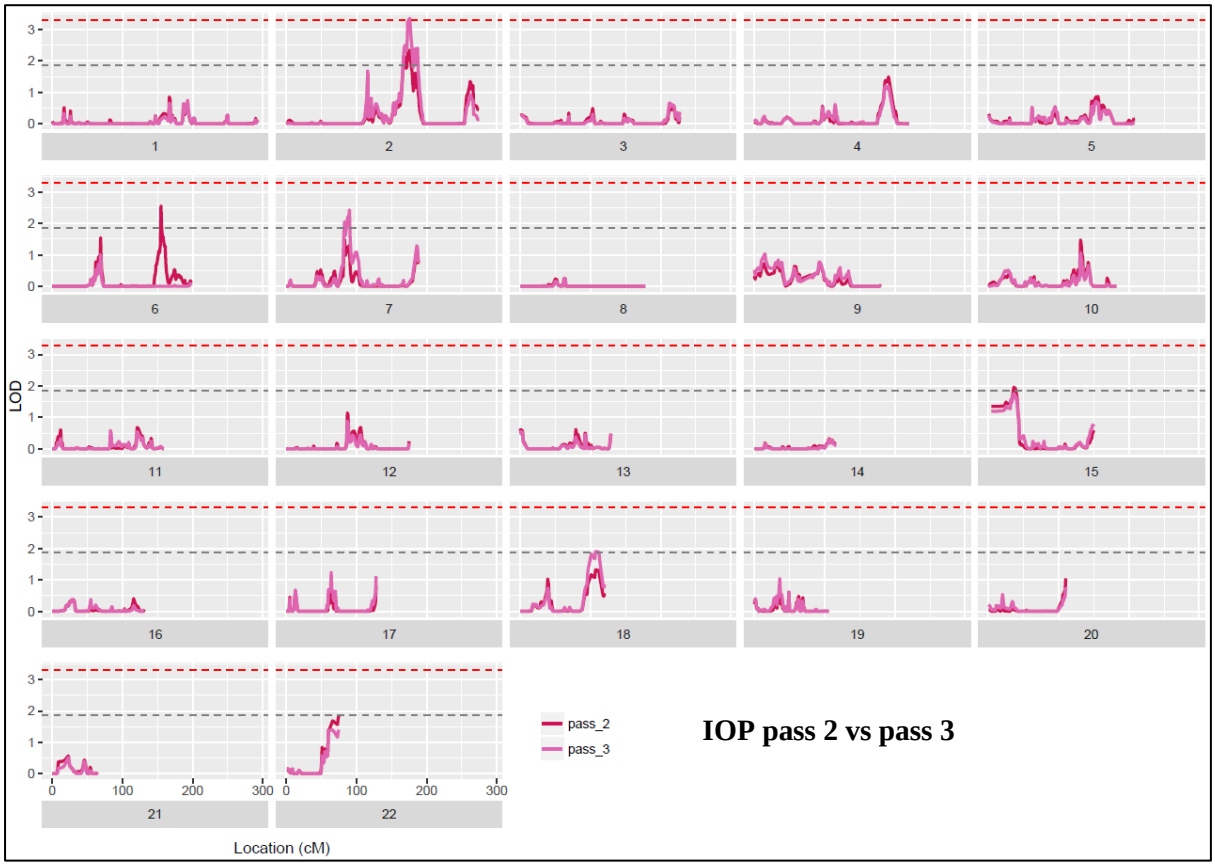
Figure 4-3 Variance components multipoint linkage analysis for chromosomes 1 to 22 for intraocular pressure traits. Intraocular pressure was unaltered **(IOP)** or medication adjusted **(IOPmed)**. **Location (cM)** = centimorgan location along each chromosome, **LOD** = logarithm of the odds, with the red dashed line indicating a significance level of 3.3 and the grey dashed line, a significance level of 1.86

The linkage plots are very similar between IOP and IOPmed, with peaks identified in the same regions (Figure 4-3). Two peaks reached full statistical significance for IOPmed, at LOD = 3.3, on chromosomes 9 and 15. These two peaks reached suggestive significance for IOP. Peaks reaching suggestive significance for both intraocular pressure traits were identified on chromosome 2, 6 and 7, with a peak on chromosome 3 only reaching suggestive significance for IOPmed. Overall, the IOPmed trait generated larger peaks for the linkage analyses. The chromosome 2 peak was the only major peak with a slightly lower maximum LOD score for the IOPmed trait (0.2 lower for IOPmed). This peak was also narrower for IOPmed, with a single peak, whereas the IOP plot shows a broader peak with additional minor peaks.

4.3.4 Multiple passes for intraocular pressure

Multiple linkage passes were conducted for IOP and IOPmed, to determine whether subsequent linkage signals were independent of previously detected QTLs. Figures 4-4 and 4-5 show the multiple passes for IOP and IOPmed respectively. Both traits fulfilled the criteria for four passes until there were no peaks reaching a maximum LOD score of 2. Both traits also followed the same order of conditioning on the QTLs, with chromosome 9 the largest peak in the first pass, followed by chromosome 6 in the second pass and chromosome 2 in the third pass. When conditioned on chromosome 7 in the fourth pass, no peaks remained with a LOD score of 2 in the final pass for either trait. Figures 4-4 and 4-5 show that conditioning on the largest QTL for each pass did not reduce the linkage signal of the remaining peaks.


LOD
Location (cM)
pass_1
pass_2
IOP pass 1 vs pass 2


LOD
Location (cM)
pass_2
pass_3
IOP pass 2 vs pass 3

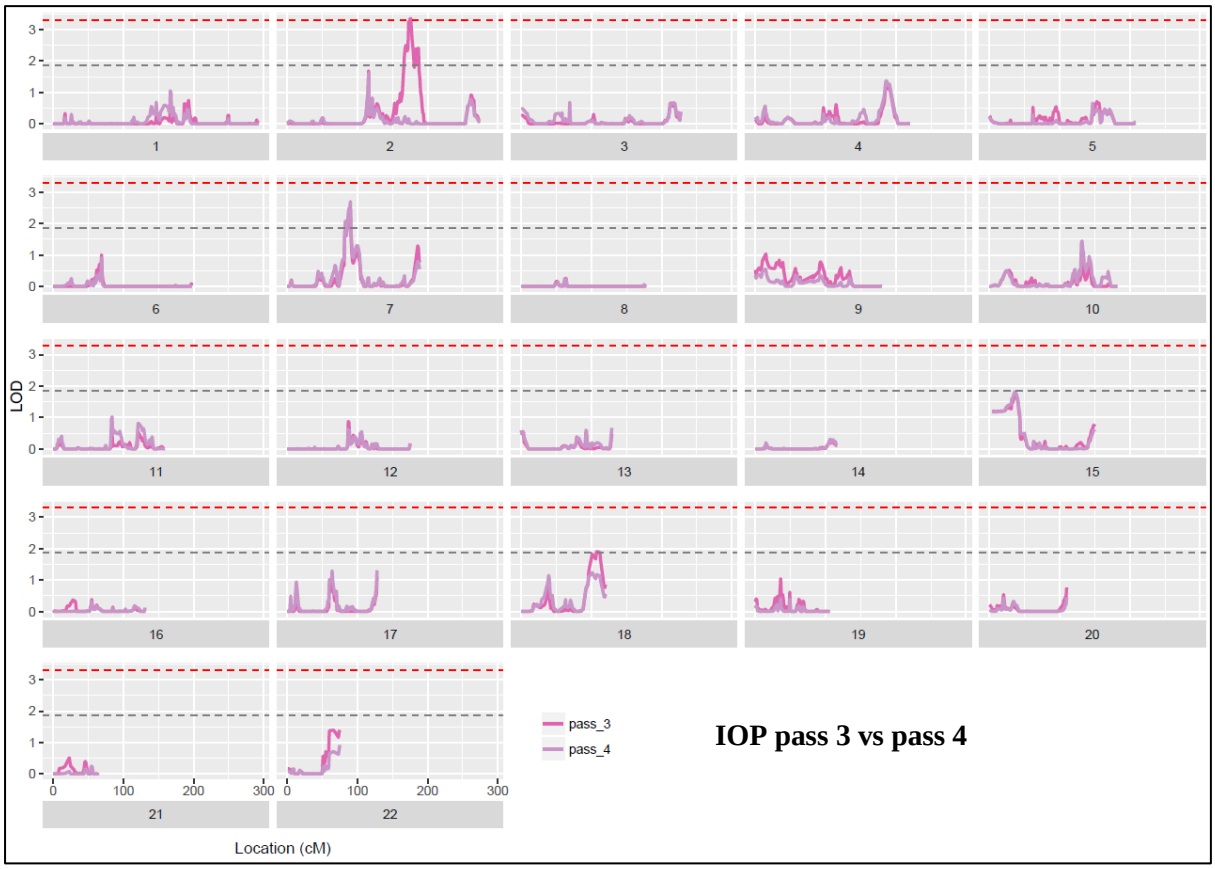
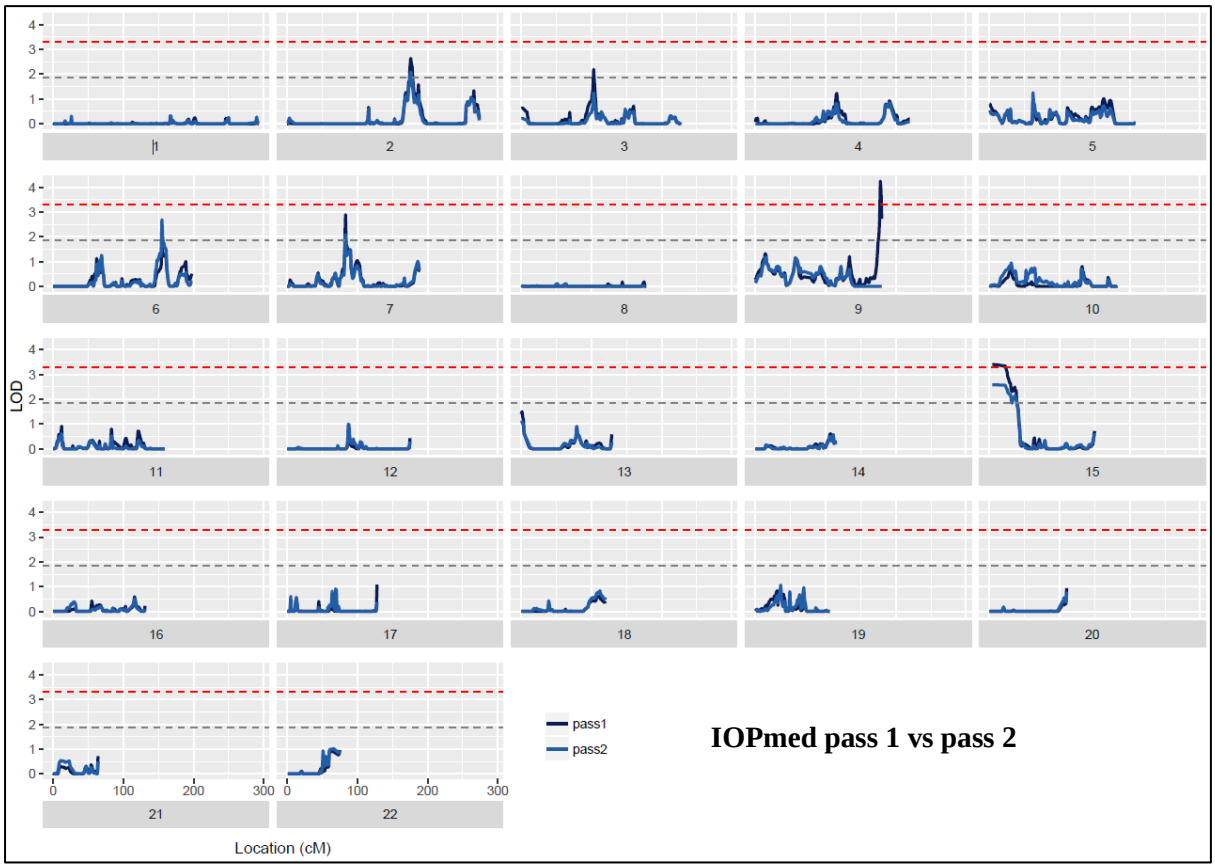
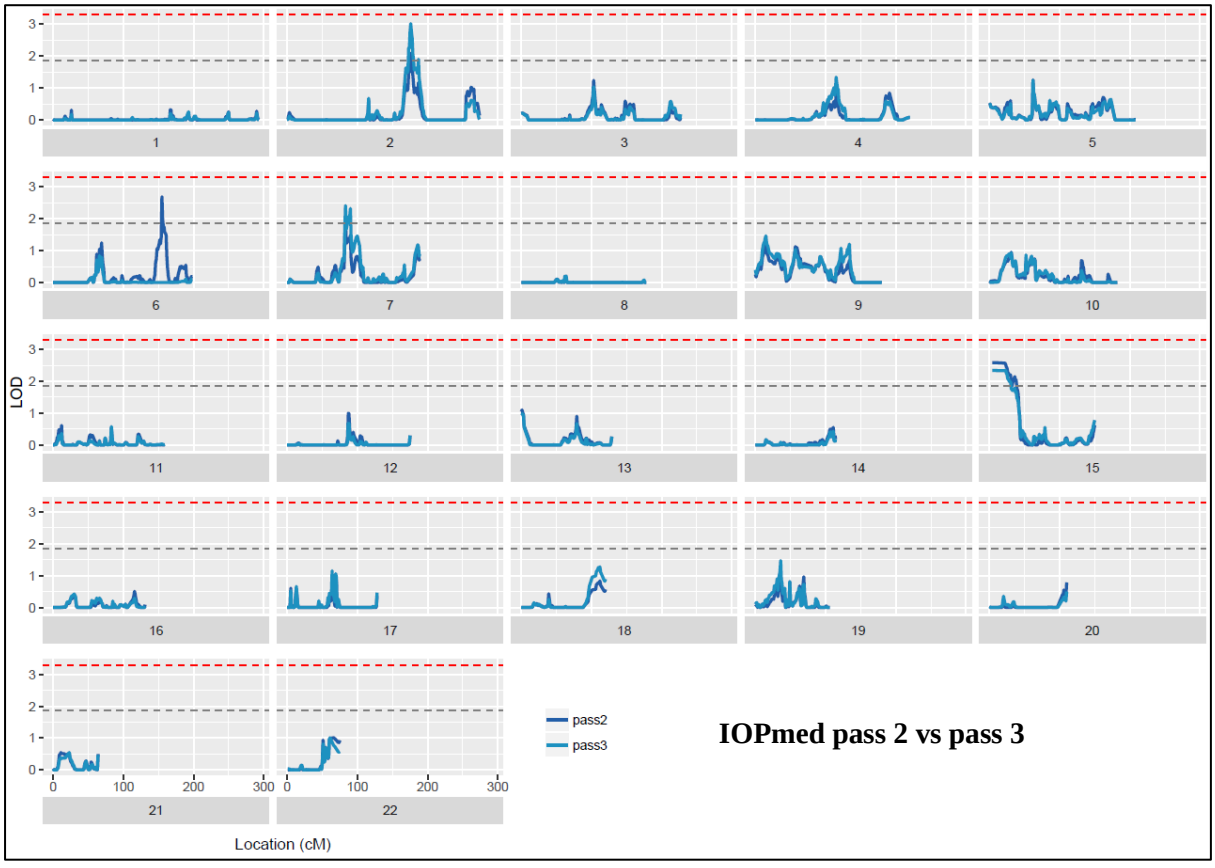


Figure 4-4 Passes 1 to 4, drawn in pairs, of variance components multipoint linkage analysis for IOP for chromosomes 1 to 22 Location (cM) = centimorgan location along each chromosome, **LOD** = logarithm of the odds, with the red dashed line indicating a significance level of 3.3 and the grey dashed line, a significance level of 1.86



IOPmed pass 1 vs pass 2



IOPmed pass 2 vs pass 3

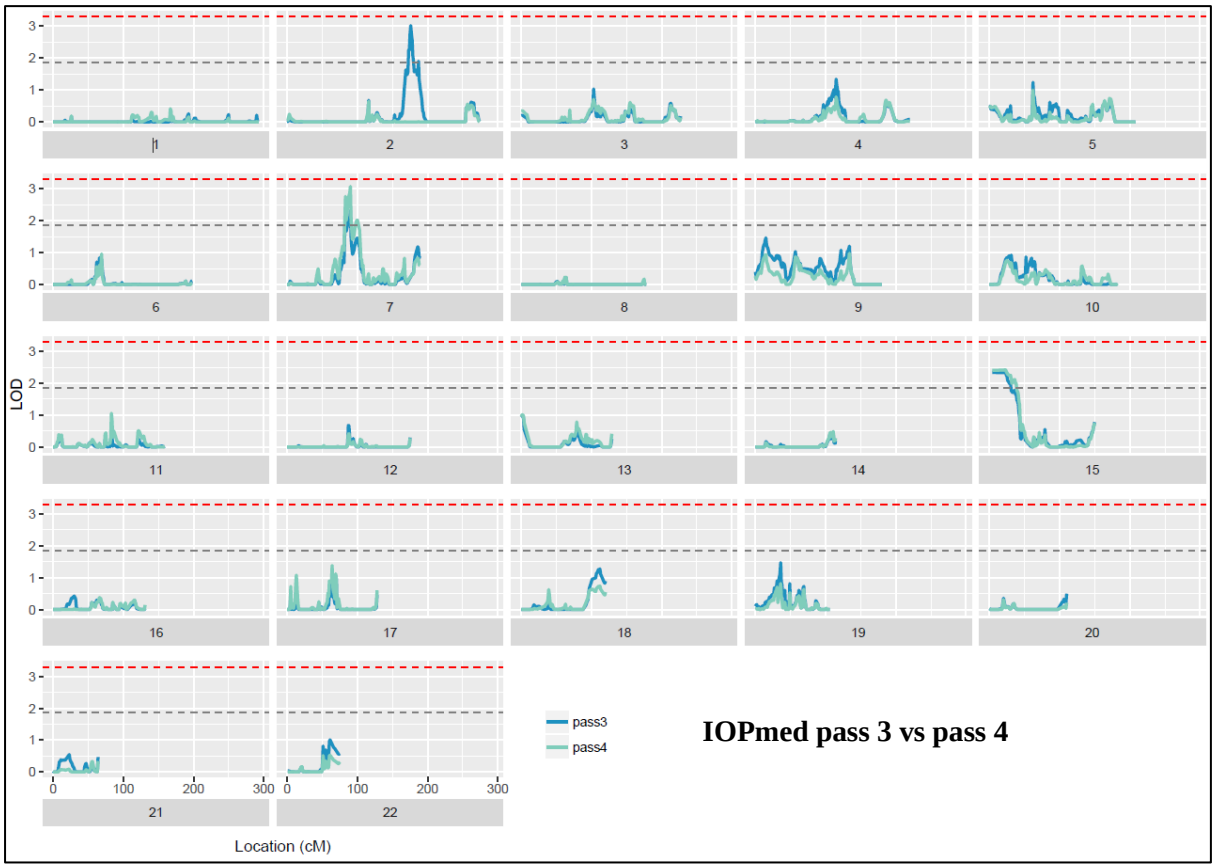


Figure 4-5 Passes 1 to 4, drawn in pairs, of variance components multipoint linkage analysis for IOPmed for chromosomes 1 to 22 Location (cM) = centimorgan location along each chromosome, **LOD** = logarithm of the odds, with the red dashed line indicating a significance level of 3.3 and the grey dashed line, a significance level of 1.86

4.3.5 Multipoint linkage analysis for VCDR

Variance components multipoint linkage analysis was also conducted on the VCDR trait (Figure 4-6). This trait was ascertainment corrected and did not require LOD adjustment. One peak, on chromosome 7, reached suggestive significance (maximum LOD = 2.5), but no QTLs were identified reaching full significance for this trait.

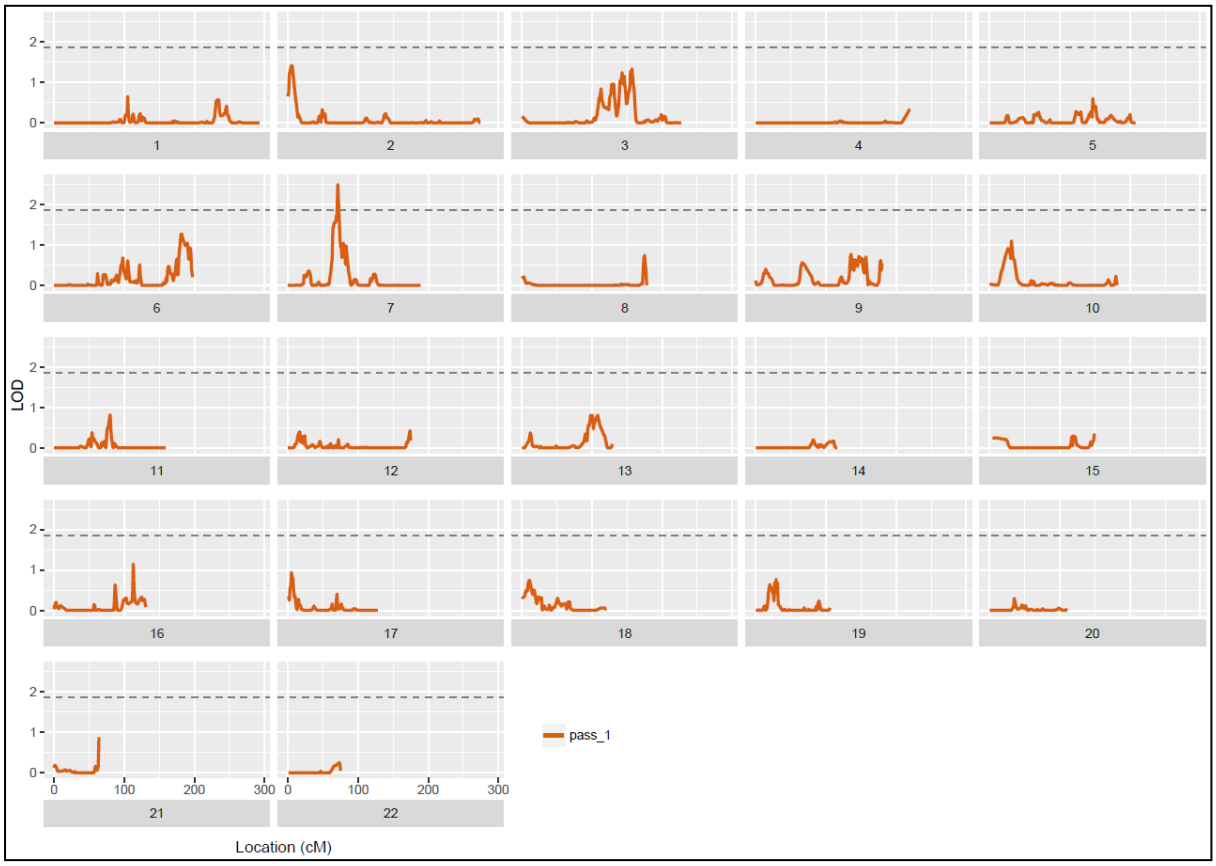


Figure 4-6 Variance components multipoint linkage analysis for chromosomes 1 to 22 for VCDR Location (cM) = centimorgan location along each chromosome, **LOD** = logarithm of the odds, with the grey dashed line indicating a significance level of 1.86

4.4 Discussion

Variance components multipoint linkage analysis of WES data was used to identify several QTLs for IOP in the five extended pedigrees of this study. Results for VCDR are not so clear cut, but possibly at least one suggestive QTL has been identified in this study. This is the first study to identify QTLs of POAG endophenotypes from WES. Although linkage analysis of quantitative endophenotypes, rather than a discrete POAG diagnosis, have proven to be successful in identifying QTLs, microsatellites and SNP genotyping panels are more commonly used to conduct the linkage analyses, with deeper sequencing conducted once QTLs have been identified [104, 133, 233]. An advantage of using WES data for the linkage analysis itself, rather than microsatellites or SNP array data, is that once a QTL has been identified, the WES data itself can be interrogated for variants with potential functional relevance. Although there are limitations to using WES in linkage studies; only a small portion of the genome is sequenced and predominantly only protein coding regions can be interrogated, there is still a great potential for identifying variants affecting disease. Although linkage analysis has been conducted from WES in other disease phenotypes [234-238], the use of WES in such large families and for endophenotypes of POAG, makes this study unique.

Using WES data for linkage analysis required optimisation at several steps to ensure accurate and appropriate data flow from one step to the next. WES data, representing the protein coding regions of the genome, is extremely dense in some areas and non-existent in other areas of the genome [239, 240], providing a new set of challenges for data preparation, analysis and interpretation. Previous studies using WES for linkage analysis pruned the variants to select common SNPs with high heterozygosity and in low LD with each other [234-238]. Our study used a different strategy, using all biallelic SNPs with MAF > 1.5% (see section 4.2.1.1), regardless of heterozygosity, and instead of LD pruning the SNPs, using the IBDLD program to condition on the existing LD to estimate IBD [152]. Other members of our group have compared the use of IBDLD for generating empirical kinships with those generated by SOLAR from reported pedigree relationships, and found them to be comparable and appropriate to use when sufficient genotype data are available [241]. The use of IBDLD to generate MIBD matrices for linkage analysis in SOLAR had also been conducted in our group, and linkage regions identified for plasma triglyceride levels [228]. These studies provided confidence that using IBDLD to estimate IBD between each of the members of the extended pedigrees and generating an empirical kinship matrix to use for the linkage analyses, was an appropriate method to use.

Heritability modelling on the endophenotypes used for linkage analysis in this study have shown that they were all significantly heritable (Table 4-3) indicating there was enough total additive genetic variance to separate into specific QTL variances, as represented by $\prod\sigma^2_{\text{qtl}}$. in equation (3) from section 4.1. The heritabilities for IOP (43%), IOPmed (50%) and VCDR (41%) in this study are well within the broad range for heritabilities for these endophenotypes which have been recently reviewed [128]. There is a great deal of variability in the literature on the heritabilities of these traits as they are dependent on the study design, sample size, method of analysis and included covariates. In our study, age was a common significant covariate to all three endophenotypes, with age^2 also significant in the pressure traits. This is unsurprising, as the risk of POAG increases with age in the general population [10, 16, 27, 29].

Two important corrections were applied to the linkage analyses in this study; ascertainment correction and empirical LOD adjustment. Ascertainment correction constrains the trait mean and standard deviation as well as those of the covariates to that of an unascertained population [167], in this case, from the Melbourne Visual Impairment Project [230]. Ascertainment correction magnifies potential differences between the families of this study, who were recruited due to enrichment of POAG, and the general population. An individual with high IOP will appear more extreme when compared to the general population than within a family which has a high average IOP. Ascertainment correction provides a more accurate representation of the linkage analyses where the families are considered not in isolation, but as part of a general population. The LOD adjustment was an important correction to make for the IOP and IOPmed linkage analyses. The correction factors which were calculated, 0.63 for IOP and 0.59 for IOPmed, suggest an inflation of LOD scores due to the non-normal distribution of these traits [212, 231]. Multiplying all LOD scores by the correction factor reduced the resultant linkage analyses by over a third, but provides confidence that peaks reaching significance are true representations of QTLs for these traits.

Our study used an adjustment for those family members who were already using pressure reducing medication prior to entering our study (see Chapter 2.5.1). A normal IOP of $\leq$ 21mmHg, in a person prescribed eye drops to reduce pressure, represents successful treatment, but this individual then is categorised as having healthy pressure, when genetically, they contribute to the group with high pressures. By applying an adjustment to IOP measurements for these individuals, a truer representation of their "unmedicated" IOP is attained, leading to a more extreme IOP phenotype measure. The adjustment used in our study was the addition of a standard deviation of the IOP pressure of all the measured family members

(SD = 6.1mmHg). This value is comparable to, and possibly underestimates, the effects of pressure reducing medication, which has demonstrated a reduction in IOP of over 8mmHg [242, 243]. Springelkamp et al. (2017), also adjusted for a "pre-medication IOP value" in their GWAS, by multiplying the IOP measurements by a factor of 1.3 [136]. It is difficult to determine which is the better method of adjustment; ours used the addition of a constant, meaning that each family member who was on pressure reducing medication prior to the start of the study had their IOP measurements increased by the same amount, regardless of their starting IOP measurement. Multiplication by a factor of 1.3 in the other study [136] increases lower pressures by a smaller amount and the larger pressures become much more inflated. Most of the medication adjusted family members in our study also had very high VCDR measures (≥ 0.8) and associated vision loss representing a more extreme phenotype.

The results presented for the unadjusted versus the adjusted pressure traits were very similar (Figure 4-3), with peaks identified in the same locations. The major difference between the two is that the IOPmed trait demonstrated enlarged peaks for chromosomes 9 and 15, reaching full statistical significance, whereas these peaks only reached suggestive significance for the IOP trait. The other difference is the chromosome 2 linkage region was narrower and tighter at the peak for IOPmed than IOP. When the multiple passes were assessed for these two traits, again they were very similar (Figures 4-4 and 4-5). The order in which the QTLs were conditioned were identical between the two traits, and more importantly, there was no loss in linkage signal of the major peaks between successive passes. This suggests that each peak is an independent QTL and not inflated due to epistatic effects. We believe that the IOPmed trait more accurately reflects the participants' IOP measures when the effects of pressure reducing medication have been considered. Also, this trait has a greater heritability (Table 4-4), with greater statistical significance, providing more power for variance components analysis. Consequently, the QTLs for the IOPmed trait was chosen for further analysis, which is discussed in Chapter 5.

A small methodological artefact, which may have been generated in the IBD estimations and/or the linkage analyses conducted in this study, was seen with the slight rises in linkage signal which occur at the end of chromosomes 12, 13, 15 17, 20, 21 and 22 (Figure 4-3). The start of the chromosomes was less affected by this artefact, although small effects on chromosomes 5 and 13 were observed. The chromosome 15 QTL, which starts at a significant LOD score for IOPmed, is unlikely to be purely artefactual, due to the difference between its size and shape compared with the equivalent position in the other chromosomes. This peak is discussed in detail in Chapter 5.

Most of the linkage regions identified in this study are novel, with one peak completely overlapping a previously identified POAG locus (Figure 4-7). The peak on chromosome 15 completely overlaps the GLC1I locus, originally identified in 15 families with an early onset form of POAG [177]. A more recent study from the same group replicated the finding at the GLC1I locus, using age of onset in an ordered subset analysis [233]. Interestingly, the linkage peak in those studies was a very similar shape to the peak in our study, with the locus occurring at the beginning of the chromosome, providing confidence that the QTL identified in our study is genuine.

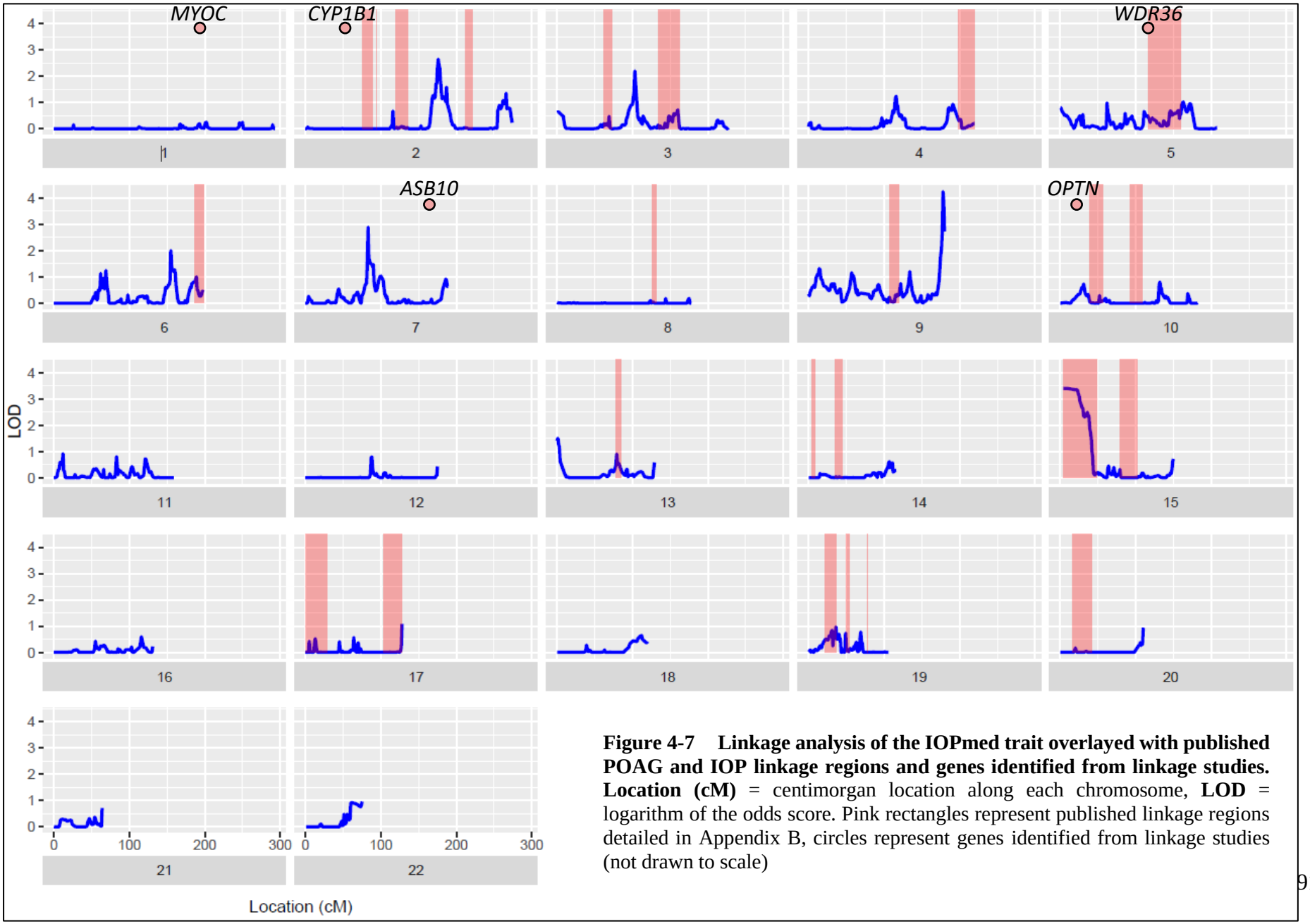


Figure 4-7 Linkage analysis of the IOPmed trait overlayed with published POAG and IOP linkage regions and genes identified from linkage studies. Location (cM) = centimorgan location along each chromosome, **LOD** = logarithm of the odds score. Pink rectangles represent published linkage regions detailed in Appendix B, circles represent genes identified from linkage studies (not drawn to scale)

The linkage analysis conducted for VCDR, which was ascertainment corrected but did not require LOD adjustment, did not identify any QTLs reaching full significance (Figure 4-6). One peak, on chromosome 7 reached suggestive significance. Although on the same chromosome as one of the intraocular pressure QTLs identified, these peaks do not overlap. This QTL is novel and does not overlap any published POAG or VCDR linkage regions. The lack of linkage signals in the VCDR analysis may be due to the nature of the VCDR measurements. These measurements only range between 0.1 and 1.0 with a VCDR of 1.0 representing advanced optic nerve damage, where the optic nerve has degenerated so much that the optic cup has enlarged to encompass the whole optic disc [14]. Although quantitative, the VCDR measurements are more categorical in nature, with increments of 0.1, and not as continuous as the IOP measurements. These VCDR ratios were multiplied by 10 to increase the standard deviation within this trait for analysis using SOLAR, but even with this adjustment, there was very little range within the data. An alternative approach would be to use SOLAR to model VCDR as a discrete trait, instead of quantitative. A threshold of ≥ 0.7 could be set for the VCDR measurements, being the threshold commonly used for POAG diagnosis [14], allowing for all family members to be classed in two groups. Although not as powerful as a quantitative linkage analysis [167, 244], this may prove to be more powerful than the analysis already undertaken. It would also be possible to conduct the linkage analysis using other programs specifically designed for categorical data. For example, the Bayesian approach of Brisbin et al. (2010), was developed for conducting linkage analysis on ordinal and categorical traits in complex pedigrees [245]. However, for the purposes of this study, the VCDR linkage analysis was left at this point, with the identification of a suggestive QTL on chromosome 7.

The use of WES data has successfully identified several QTLs for IOP in the five large, extended families of this study. Further work on identifying each family's contribution to each of these QTLs as well as finding variants with potential functional relevance within these QTLs is described in Chapter 5.

Chapter 5
Identification of variants within linkage peaks and proposal of novel IOP genes

5.1 Introduction

In the previous chapter, quantitative trait loci (QTLs) for intraocular pressure (IOP) were identified in five families. These peaks reduced the genome-wide search space to provide the opportunity to find variants and genes involved in IOP regulation, which may also be involved in POAG pathogenesis in these families. Although previous linkage analyses have identified over 20 regions associated with glaucoma or its endophenotypes [13, 27], the causative variants or genes have not been identified in most of these regions. These linkage analyses were conducted using microsatellite or SNP array markers resulting in the identification of large regions, but the genes within these regions could only be proposed as possible candidate genes. Deeper sequencing was required to determine if the candidate genes carried deleterious, potentially disease causing variants. Before the advent of next generation DNA sequencing technologies, identifying these variants from within the many genes within linkage regions was difficult, time consuming and expensive. It often took years to identify the glaucoma causing genes; *MYOC* and *OPTN* were reported as the causative genes four years after their linkage regions were published [46, 55, 59, 70] and *ASB10* was not identified until 13 years after the linkage region was published [71, 72]. Most of the linkage regions for glaucoma have not had causative genes identified at all.

Our study has identified six QTLs for IOP (Chapter 4) and the next stage was to interrogate these regions and identify variants and genes which may be important in IOP regulation. Exome sequencing data was used directly to conduct the linkage analysis. The advantage this provided was to enable the identified QTLs to be examined for functional exonic variants, without needing to conduct additional sequencing. Selected variants could then be included in the linkage model, as covariates, to determine their effects on the linkage peaks. Almasy and Blangero (2004) [246], describe this linkage conditional on measured genotype approach. Variants to be tested are coded as measured genotypes, with the dosage of the variant (0,1 or 2) in each individual used as the covariate, similar to that used in MGA (Chapter 2.6.3).

Linkage can then be conducted conditional on this measured genotype, which is included as a fixed effect on the trait mean and not a component of the trait variance [167]. The effect of the measured genotype reduces the variance available for the linkage component of the analysis. If the measured genotype accounts for much of the fixed effects of the mean, less variance is available for the linkage component and the LOD score of the linkage region is markedly reduced. If there are several variants affecting a peak, the measured genotype of one variant may not be sufficient to reduce the variance available and consequently the peak may not be affected, or only reduced slightly. It should be noted that this method is not able to determine whether the variant being tested is a variant with potential functional relevance, or whether it is on the same haplotype as the functional variant [246].

Prioritised variants were then assessed using *in silico* tools to predict which genes might be involved with IOP regulation in the families of this study. Complex diseases are likely to result from involvement of both common and rare variants [119]. The families used in this study, which are enriched for POAG, are ideal to use to search for rare variants influencing this complex disease, due to the potential of enrichment of rare variants as they segregate through the generations. Using families ascertained for a disease as well as using endophenotypes of the disease both increase the power to identify functional variants [211]. To identify variants influencing IOP, we selected the linkage regions identified in Chapter 4 using the IOPmed trait, which provided a more realistic representation of IOP measures of the family members.

5.1.1 Aim of this study

To propose candidate genes which may be involved in IOP regulation in the families of this study.

Specifically, this study will use *in silico* tools to investigate genetic variants within the most significant peaks identified from the linkage analysis in Chapter 4, and propose genes which may be involved in IOP regulation.

5.2 Methods

Linkage peaks identified for the IOPmed trait, in Chapter 4, were selected for further investigation in this study if they reached at least suggestive significance at a LOD of 1.86 based on the Lander and Kruglyak significance criteria for linkage proposed in 1995 [232]. Linkage peaks on chromosomes; 2, 3, 6, 7, 9 and 15 fulfilled this criterion, with chromosomes 9 and 15 peaks also being significant at the genome-wide level (LOD ≥ 3.3).

5.2.1 Defining linkage intervals

For each of the linkage peaks reaching a LOD score of at least 1.86, the boundaries of the peak interval were defined as the cM position where the LOD dropped to 0.59; which is equivalent to a nominal p-value of 0.05 [232, 247]. All variants within the boundaries were then extracted from the original bcbio called VCF file, to ensure all of the sequenced variants were included in the linkage intervals. This VCF file was generated prior to Mendelian error checking (described in Chapter 4.2.1.2), hence some of these variants may have had errors. The full VCF file was intentionally not checked with PedCheck, as we did not wish to discount any variants which may have included a minor sequencing or calling error. However, all variants of interest were checked manually to ensure there were no Mendelian errors. These variants were annotated using ANNOVAR as described in Chapter 2.4.2.

5.2.2 Family contribution to linkage peaks

The first stage in narrowing down the number of variants for further analysis was to determine the contribution of each family to the linkage peaks. Variants found predominantly within those families were then prioritised further. Based on the number of individuals and the trait heritability there was insufficient power to analyse pedigrees individually. Accordingly, to retain sufficient heritability to conduct the analyses, the per-family contributions were determined by conducting linkage analysis on four families at a time, with a single family excluded for each analysis. A marked reduction of the linkage peak when a family was excluded is indicative of the contribution of that family to the IOPmed linkage at that peak. The linkage analysis protocol is described fully in Chapter 4.2.2.

5.2.3 Measured genotype association analysis

MGA was conducted for the IOPmed trait on the variants within the intervals for each of the linkage regions. MGA, a family-based association test, is described fully in Chapter 2.6.3. A maximum of 10% of low confidence and missing calls per variant (as defined in Chapter 2.4.2) was required for inclusion in the MGA analysis.

5.2.4 Linkage conditional on measured genotype

Individual variants were included in the linkage model as a covariate, prior to ascertainment correction, to determine their contribution to the linkage peak, as described by Almasy and Blangero, (2004) [246]. Variants were added individually, or in combination with other variants. The genotypes of each individual, for the variant of interest, were scored as a dosage, based on the number of copies present; with 0 for non-carriers, 1 for heterozygous carriers and 2 for homozygous variant carriers. Linkage conditional on the measured genotype of the variant of interest was then conducted. All the linkage analyses were conducted on the IOPmed trait, with both an ascertainment correction and a LOD adjustment applied (see Chapter 4.2.2 and 4.2.3)

5.2.5 Prioritisation of variants for further analysis

Variants identified from within the linkage intervals on each chromosome were prioritised based on the following criteria.

5.2.5.1 Presence in families contributing to the linkage peak

Variants were prioritised if they were present predominantly in the family or families contributing to the linkage peak (see section 5.2.2). As this study has focussed on rare variants, there was an expectation that the same rare variant would not be identified in multiple unrelated families. For each variant, at least 90% of the calls must have been identified in one of the families contributing to the peak. A strict 100% threshold was not chosen to allow for possible sequencing or calling errors, which may occur with joint-calling across multiple families, so as not to discount an identified variant and also to allow for uncommon rather than only rare or private variants.

5.2.5.2 MGA results

Variants were prioritised if p-values were less than or equal to 0.05 in the MGA analysis (see section 5.2.3).

5.2.5.3 Reduction in LOD score

Variants fulfilling both criteria above were tested for linkage conditional on their measured genotype (see section 5.2.4). Variants were prioritised if they reduced the LOD score of the linkage peak by a minimum of 1 LOD unit.

5.2.6 Additional variants

Variants were also assessed if they were in the majority of the same individuals as those already prioritised. The purpose of assessing these variants was to determine if the haplotype encompassing prioritised variants could also be harbouring other, potentially deleterious, variants in linkage.

5.2.7 Further analysis of prioritised variants and genes

Prioritised and additional variants were assessed further for potential functional affects at both the variant and gene level.

5.2.7.1 Functional annotations and conservation

The UCSC Genome Browser (http://genome.ucsc.edu/cgi-bin/hgGateway) was used to determine which mRNA transcripts, eQTL sites and miRNA binding sites (for 3' UTR variants) and active promoter regions (for 5' UTR variants) overlap with the variant location. Conservation of variant sites was assessed with ANNOVAR annotations (Chapter 2.4.2) for Genomic Evolutionary Rate Profiling (GERP) and followed with UCSC's GERP track for any missing annotations. The Human Splicing Finder v3.1 (http://www.umd.be/HSF/) was used to assess whether the identified variants could potentially affect splicing of mRNA transcripts.

5.2.7.2 Gene expression in ocular tissues

Two ocular tissue datasets were used to assess the expression of prioritised genes in ocular tissues. We accessed a gene expression dataset (provided by Prof. Jamie Craig and Dr Tiger

Zhou, Flinders University) generated from human ocular tissues (referred to as "in-house RNAseq data"). The method for generating this RNAseq data is described in detail by Macgregor et al, (2018) [248]. Briefly, RNA was extracted from 48 samples obtained from the donated eyes of 16 individuals. Nine distinct ocular tissues were used, including the trabecular meshwork and ciliary body. Sequencing was undertaken using an Illumina platform and mapped to the hg19 human reference genome. Normalised counts per million (CPM) data were collected from the transcripts of 21,962 RefSeq protein coding genes.

Gene expression data for the same ocular tissues was also accessed from Ocular Tissue Database (OTD, https://genome.uiowa.edu/otdb/), which uses data generated from the pooled samples from six individuals. The OTD uses a normalised "PLIER" score to display gene expression results obtained from microarray data [249].

5.2.7.3 Gene interaction analysis

To determine whether there was evidence for known (published) gene/gene product interactions between the prioritised genes and previously identified IOP or POAG associated genes, we used the custom pathway generation functionality in Ingenuity Pathway Analysis (IPA, Winter 2019 Release, QIAGEN Inc) [250]. The IPA knowledge base was used to identify any published interactions between prioritised genes from this current study and a combined set of 129 previously identified IOP genes from Danford et al. (2017) [251] and POAG genes identified in Janssen et al. (2013) [13].

5.2.7.4 Significance of prioritised regions in GWAS data

Prioritised regions from this study were tested for significance in the recent IOP and glaucoma meta-analyses conducted by MacGregor et al. (2018) [248]. These meta-analyses were published after the literature review was completed for our current study (see Chapter 3.3.1) and as such, those findings were not included in the POAG database (Appendix B). We accessed the GWAS summary statistics (provided by Prof. Stuart MacGregor, Queensland Institute of Medical Research) to assess our prioritised variants in a large population of individuals including glaucoma cases. The IOP GWAS was conducted on participants from the UK Biobank (n=103,914) and the International Glaucoma Genetic Consortium (n=29,578). Regions from our study were also tested for significance in glaucoma cases (n=11,018) and controls (n=126,069), from the UK Biobank and the Australian and New Zealand Registry of Advanced Glaucoma. This was a combined glaucoma dataset which included, but was not

limited to POAG cases. LocusZoom (http://locuszoom.org/) was used to visualise the GWAS data for each of the prioritised regions identified in this current study. The 1000 Genomes November 2014 European database was used, within LocusZoom, for the LD population in these plots.

5.3 Results

5.3.1 Linkage intervals

As discussed in Chapter 4 there were six linkage peaks for IOPmed that exceeded the LOD threshold for suggestive linkage and were therefore selected for further investigation. Table 5-1 shows the identified linkage peaks, their maximum LOD scores and the positions of the intervals used to select variants for evaluation.

Table 5-1 IOPmed linkage peaks and linkage intervals for analysis of variants within peaks

Linkage peaks				Linkage intervals		
Maximum LOD score	Chr	Position of the maximum LOD (cM)	Position of the maximum LOD (bp)	Position (cM)	Position (bp)	Cytogenetic location
4.25	9	178	139,217,089	172 180	137,593,099 141,111,553	9q34.3
3.40	15	4	20,876,518	3 43	20,743,796 31,197,564	15q11.2-13.2
2.89	7	83	63,674,464	79 93	53,103,371 77,756,724	7p11.2-q11.23
2.64	2	176	158,384,798	167 192	147,345,425 174,774,366	2q23.1-31.1
2.19	3	103	75,470,807	95 107	71,408,319 86,116,240	3p13-12.1
2.00	6	155	146,120,584	151 163	142,539,730 151,859,314	6q24.2-25.1

Chr = chromosome, **cM** = centimorgan, **bp** = base pair

5.3.2 Family contribution to linkage peaks

Table 5-2 shows the effects of excluding individual families, from the linkage analysis, on the maximum LOD score for each of the linkage peaks for the IOPmed trait. Figure 5-1 displays the effects of family exclusion across the whole chromosome for each of the chromosomes identified with linkage peaks. Four out of the five families contributed to the six linkage peaks.

GTAS04 was the only family whose exclusion from the linkage analysis did not reduce any of the linkage peaks. The whole chromosome 7 peak was dramatically reduced when a single family, 98002, was excluded from the linkage analysis (Figure 5-1d). Conversely, the peak on chromosome 6 had minor contributions from four families (Figure 5-1c). Each of the four families was the major contributor to at least one linkage peak; Family 93001 to chromosome 9 (Figure 5-1e), Family 95002 to chromosome 15 (Figure 5-1f), Family 98002 to chromosomes 2 and 7 (Figure 5-1a and d) and GTAS54 to chromosome 3 (Figure 5-1b).

Table 5-2 Family contributions to linkage peaks. LOD scores for the 5 families together and for each family excluded individually. Families were included in the table if their exclusion from the linkage analysis resulted in a reduction in maximum LOD score. Numbers in bold represent the major family's contribution to each peak.

5 families maximum LOD	Chr	Maximum LOD when family was excluded and (reduction in maximum LOD)			
		93001	95002	98002	GTAS54
4.25	9	**1.64 (-2.61)**		2.43 (-1.82)	
3.40	15	2.93 (-0.47)	**1.41 (-1.99)**		2.54 (-0.86)
2.89	7			**0.04 (-2.85)**	
2.64	2		1.73 (-0.91)	**1.50 (-1.14)**	1.91 (-0.73)
2.19	3	0.98 (-1.21)	1.51 (-0.68)		**0.83 (-1.36)**
2.00	6	1.69 (-0.31)	1.04 (-0.96)	1.52 (-0.48)	1.33 (-0.67)

Chr = chromosome, **LOD** = logarithm of the odds

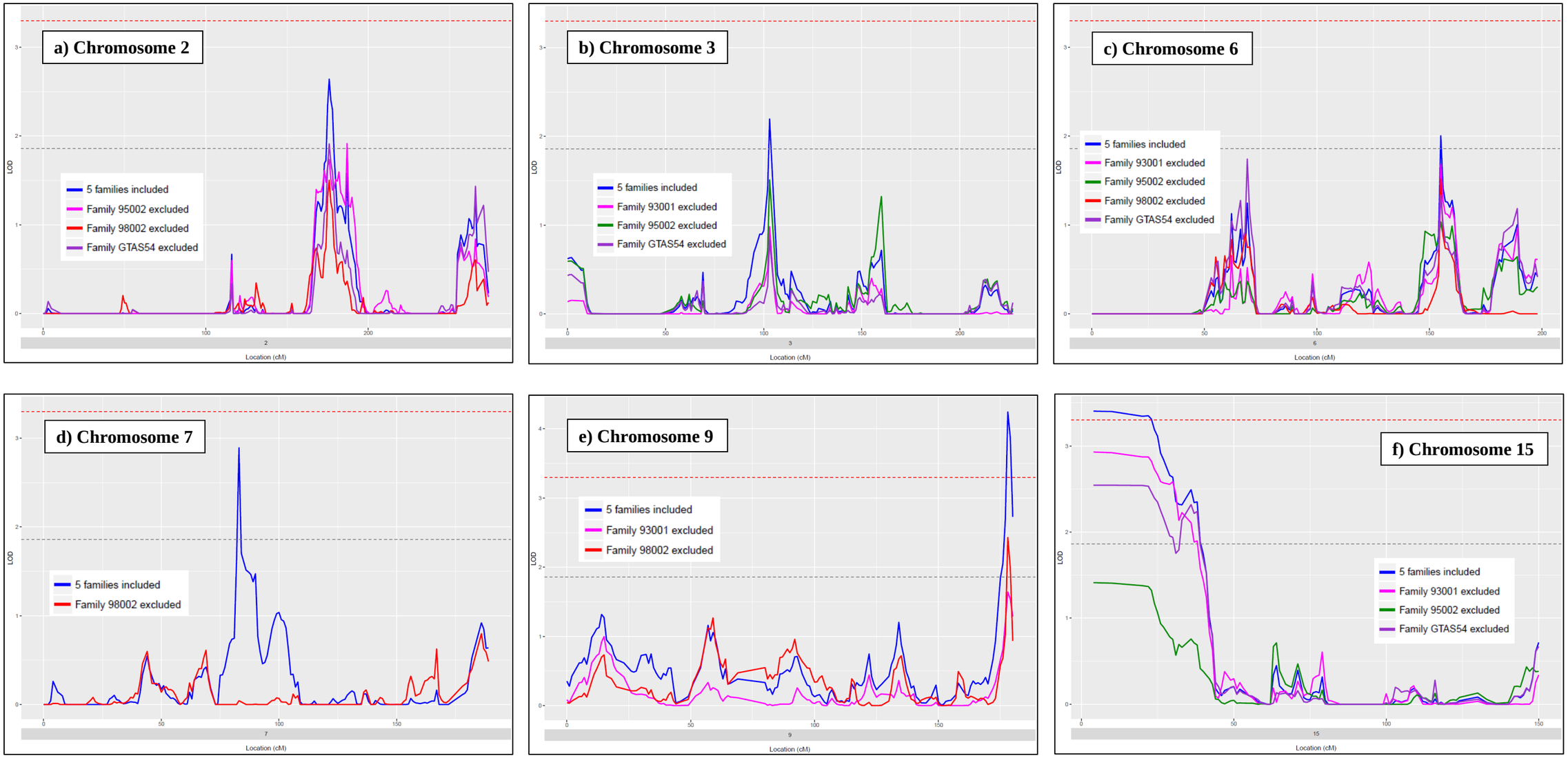


Figure 5-1 Effect of excluding individual families on the six linkage peaks for IOPmed. Individual families are only shown if there was a reduction in peak LOD score when that family was excluded from the linkage analysis. Chromosomes: 2 **a)**, 3 **b)**, 6 **c)**, 7 **d)**, 9 **e)**, and 15 **f)**, five families together = blue, exclusion of individual families: 93001 = pink, 95002 = green, 98002 = red, GTAS54 = purple. **Location (cM)** = centimorgan location along each chromosome, **LOD** = logarithm of the odds, with the red dashed line indicating a significance level of 3.3 and the grey dashed line, a significance level of 1.86

5.3.3 Prioritised variants across all linkage intervals

Variants prioritised (using the criteria in section 5.2.5) across all of the linkage intervals are shown in Table 5-3. Appendix E displays all of the variants with MGA results of $p \leq 0.05$, and indicates those found predominantly in a single family. These variants were tested for linkage conditional on their measured genotype. The chromosome 6 linkage interval was the only one in which a variant fulfilling all of the prioritisation criteria could not be identified. Prioritised variants per linkage interval ranged from a single variant (for chromosome 15) through to five for chromosome 2. Merged cells in Table 5-3 represent variants identified in the same individuals. As the MGA was conducted using a dosage of the minor allele as a covariate, it is impossible to distinguish between the effects of the variants occurring in the same individuals, including variants identified in different genes. Variants identified in identical family members were found in the chromosome 2, 7 and 9 linkage intervals. The MGA analysis requires both phenotypic and genotypic data, and those individuals without IOP measurements were not included in this analysis. This is demonstrated in Table 5-3, with differing minor allele copy (MAC) numbers in the MGA analysis and in the family specific MAC columns. One homozygous carrier of the *RIF1* variants and a heterozygous carrier of the *CALN1* variants, both from family 98002, did not have IOP measurements and were not included in the MGA analysis.

.

Table 5-3 Prioritised variants across all linkage intervals

										MGA				Family specific MAC			
Chr	**bp position (hg19)**	**Variant ID**	**Gene**	**Gene region**	**Ref**	**Alt**	**CADD**	**gnomAD genome NFE**	**AF**	**p-value**	**beta-value**	**variance explained (%)**	**MAC**	**93001**	**95002**	**98002**	**GTAS54**
2	152,317,669	rs116443402	*RIF1*	exonic	G	A	0.3	0.020	0.072	0.0215	1.71	1.74	34			36	
2	152,331,995	rs372957656	*RIF1*	3' UTR	GTTG	-	1.6	0.020									
2	158,177,675	rs749173309	*ERMN*	3' UTR	C	A	3.5	< 0.001	0.012	0.0012	7.99	3.27	6		6		
2	166,810,232	rs572065013	*TTC21B*	5' UTR	G	A	6.4	0.004	0.018	0.0079	5.02	3.96	9				9
2	172,953,525	rs186854626	*DLX1*	3' UTR	C	T	9.9	0.002									
3	73,437,185	rs61732653	*PDZRN3*	exonic	A	G	8.7	0.012	0.020	0.0117	4.47	3.59	10			1	9
3	75,471,002	rs184945375	FAM86DP	ncRNA	C	T	3.4	0.004	0.020	0.0190	4.24	2.82	10				10
7	71,249,529	-	*CALN1*	3' UTR	G	C	1.7	*	0.028	0.0148	4.28	2.86	13			14	
7	71,249,770	-	*CALN1*	3' UTR	T	-	3.3	< 0.001									
9	139,296,703	-	*ENTR1*	3' UTR	G	A	1.3	*	0.030	0.0066	5.18	4.91	15	15			
9	140,342,693	rs746921357	*NSMF*	3'UTR	G	A	10.0	0.001									
15	28,474,713	-	*HERC2*	exonic	C	T	10.0	*	0.010	0.0005	10.34	4.58	5		5		

Chr = chromosome, **bp** = base pair
3' UTR - 3' untranslated region, **5' UTR** - 5' untranslated region, **exonic** = protein coding region (synonymous)
Ref = reference allele, **Alt** = alternate allele
CADD = combined annotation dependent depletion phred scores (version 1.4)
gnomAD genome NFE = variant allele frequency in non-Finnish Europeans in the gnomAD genome database. ***** = variant not found in this database
AF = frequency of the variant allele in the 5 families
MGA = measured genotype association analysis
beta-value = effect size per allele for IOP, measured in mmHg,
MAC = minor allele copies as used for MGA analysis.
Family specific MAC = minor allele copies in each family. May differ from MAC used for MGA if IOP information was not also available

Of the prioritised variants in Table 5-3, none were non-synonymous and were three synonymous protein coding variants; in *RIF1*, *PDZRN3* and *HERC2*. Most were 3' UTR variants, including the two deletions. None of the 3' UTR variants overlapped identified miRNA sites or eQTL sites. The one 5' UTR variant, in *TTC21B*, lies within the active promoter of this gene within a region of multiple transcription factor binding sites. Most of the prioritised variants had low CADD scores, with only the *NSMF* and *HERC2* SNPs reaching scores in the top 10% for predicted deleteriousness. Variants other than those in *RIF1* were much rarer in the gnomAD database, observed at or below 1% in the European population. Several variants were novel or observed at a frequency of below 0.1% in this database. The frequency of these variants was much higher in the five families, ranging from 1% for the *HERC2* variant to over 7% for the *RIF1* variants. All of the variants in Table 5-5 were associated with an increase in IOP, ranging from an average 1.7mmHg for each allele of the *RIF1* variants, through to over 10mmHg per allele for the *HERC2* variant. In fact, all of the variants, other than the *RIF1* variants, were associated with a minimum of 4mmHg increase in IOP, per allele. The contribution to the IOPmed trait variance, per allele, ranged from just under 2% to nearly 5%. This variance explained was not correlated with the number of alleles present in the families; the *HERC2* variant found only in five individuals contributed to the same amount of IOP variance as the *ENTR1* variant and nearly three times the variance of the *RIF1* variants, found in many more family members. Of all the variants in Table 5-3, the *HERC2* variant was found in the fewest individuals, but was associated with the largest average increase in IOP.

5.3.3.1 Evolutionary conservation and predicted splicing effects of variants

Table 5-4 shows GERP conservation scores and predicted splicing effects of both the prioritised (section 5.2.5) and additional (section 5.2.6) variants. Positive GERP values predict more evolutionarily conserved base pair positions, with maximum score possible of 6.18, and negative values predict the more neutral positions. Zero GERP scores represent a non-meaningful estimate due to poor alignment. The results displayed in this table are detailed in the relevant sections on the linkage regions below.

Table 5-4 Conservation scores and predicted splicing effects of prioritised and additional variants identified from linkage regions

Chr	bp position (hg19)	Variant ID	Gene	Gene region	Ref	Alt	GERP score	Potential splicing alteration
2	152,317,669	rs116443402	*RIF1*	exonic	G	A	-0.1	no
2	152,331,995	rs372957656	*RIF1*	3' UTR	GTTG	-	0.0	New ESS site
2	152,417,767	rs115631125*	*NEB*	exonic	C	A	-10.9	ESE site broken
2	153,378,462	rs138926754*	*FMNL2*	exonic	T	C	-2.6	ESE site broken
2	158,177,675	rs749173309	*ERMN*	3' UTR	C	A	3.3	ESE site broken
2	159,536,990	rs148782148*	*PKP4*	# exonic p.(Asp > Val)	A	T	5.5	ESE site broken New ESS site
2	166,810,232	rs572065013	*TTC21B*	5' UTR	G	A	-4.9	no
2	172,953,525	rs186854626	*DLX1*	3' UTR	C	T	3.7	ESE site broken New ESS site
3	73,437,185	rs61732653	*PDZRN3*	exonic	A	G	-8.9	ESE site broken
3	75,471,002	rs184945375	FAM86DP	ncRNA	C	T	-2.3	ESE site broken
7	71,249,529	-	*CALN1*	3' UTR	G	C	1.7	ESE site broken
7	71,249,770	-	*CALN1*	3' UTR	T	-	2.7	ESE site broken
9	139,296,703	-	*ENTR1*	3' UTR	G	A	-2.7	New ESS site
9	140,342,693	rs746921357	*NSMF*	3'UTR	G	A	-0.3	ESE site broken
15	28,474,713	-	*HERC2*	exonic	C	T	0.2	no

*** =** additional variants (from section 5.2.6) which are discussed in the relevant section on their loci below.
Chr = chromosome, **bp** = base pair
3' UTR - 3' untranslated region, **5' UTR** - 5' untranslated region, **exonic** = protein coding region (synonymous unless otherwise indicated), **#** = non-synonymous variant
Ref = reference allele, **Alt** = alternate allele
GERP = Genomic Evolutionary Rate Profiling score
Potential splicing alteration = as identified using the Human Splicing Finder v3.1
ESE = exonic splicing enhancer, **ESS** = exonic splicing silencer

5.3.3.2 Gene expression in ocular tissues

Expression of the genes identified from prioritised (section 5.2.5) and additional (section 5.2.6) variants in the ocular tissue database (OTD) and our in-house RNAseq dataset are shown in Table 5-5. These gene expression results are discussed in the relevant sections on the linkage regions below.

Table 5-5 Gene expression of prioritised and additional genes in IOP relevant ocular tissues

			Ocular Tissue Database (PLIER score)		In-house RNAseq (counts per million)	
Chr	**Gene**	**HGNC ID**	**CB**	**TM**	**CB**	**TM**
2	***RIF1***	23207	40.3800	59.2821	11.7138	38.0609
2	***NEB****	7720	25.8328	128.7910	0.4436	44.0060
2	***FMNL2****	18267	64.6647	77.1831	16.2269	48.3394
2	***ERMN***	29208	--	--	0.4959	1.1282
2	***PKP4****	9026	158.2350	115.8910	184.2291	47.4201
2	***TTC21B***	25660	34.5452	58.7574	8.0150	18.8125
2	***DLX1***	2914	40.4993	34.8669	0.4999	0.6300
3	***PDZRN3***	17704	51.2939	26.7320	74.1244	88.6112
3	***FAM86DP***	32659	22.5021	10.7506	5.2495	5.9587
7	***CALN1***	13248	22.9616	16.9850	0.8475	0.4130
9	***ENTR1***	10667	58.9830	41.1306	24.7005	23.3657
9	***NSMF***	29843	--	--	75.5354	127.2043
15	***HERC2***	4868	33.2641	48.7907	87.7595	103.2685

*** =** additional genes (from section 5.2.6) which are discussed in the relevant section on their loci below.
Chr = chromosome, **HGNC ID** = HUGO gene nomenclature committee identification number
CB = ciliary body, **TM** = trabecular meshwork, **--** = gene not assessed in the ocular tissue database

5.3.4 Analysis of individual linkage intervals

Each of the linkage intervals is discussed in detail below, in the order of descending maximum LOD score of the peaks (as shown in Table 5-1).

5.3.4.1 Chromosome 9q34.3 linkage interval

The chromosome 9 linkage peak was the most significant of all six IOP linkage peaks identified and reached full statistical significance with a LOD score of 4.25. Table 5-2 and Figure 5-1e both indicate that family 93001 contributed most to this peak with family 98002 making a minor contribution. Two variants fulfilled the criteria for prioritisation; in *ENTR1* (novel variant located at chr9:139296703) and *NSMF* (rs746921357), which were identified in the same 15 individuals from family 93001 (Table 5-3). Inclusion of the measured genotype of these variants in the linkage model reduced the maximum LOD score by over 1, but did not account for the majority of the peak (Figure 5-2). As with the MGA results, variants identified in the same individuals result in identical linkage peaks, thus one plot represents both variants. Both of these variants were in the 3' UTR regions of their genes. Both were rare variants, with the *ENTR1* variant not observed in the gnomAD database and the *NSMF* variant with a minor allele frequency of 0.001. These variants were associated with over 5mmHg average IOP increase and contributed to nearly 5% of the trait variance. Both of these variants were predicted to affect splicing but neither of them are at evolutionarily conserved sites (Table 5-4). *NSMF* was not included in the OTD, but showed high expression in our in-house RNAseq dataset, in both the ciliary body and trabecular meshwork (Table 5-5). In fact, *NSMF* had the highest expression in the trabecular meshwork of all the assessed genes. *ENTR1* was expressed at lower levels than *NSMF* in our RNAseq dataset and showed expression in the OTD at a similar level to other assessed genes. Neither of these genes were found to interact with any of the published genes assessed in the IPA analysis and did not reach significance in the published IOP and glaucoma meta-analyses.

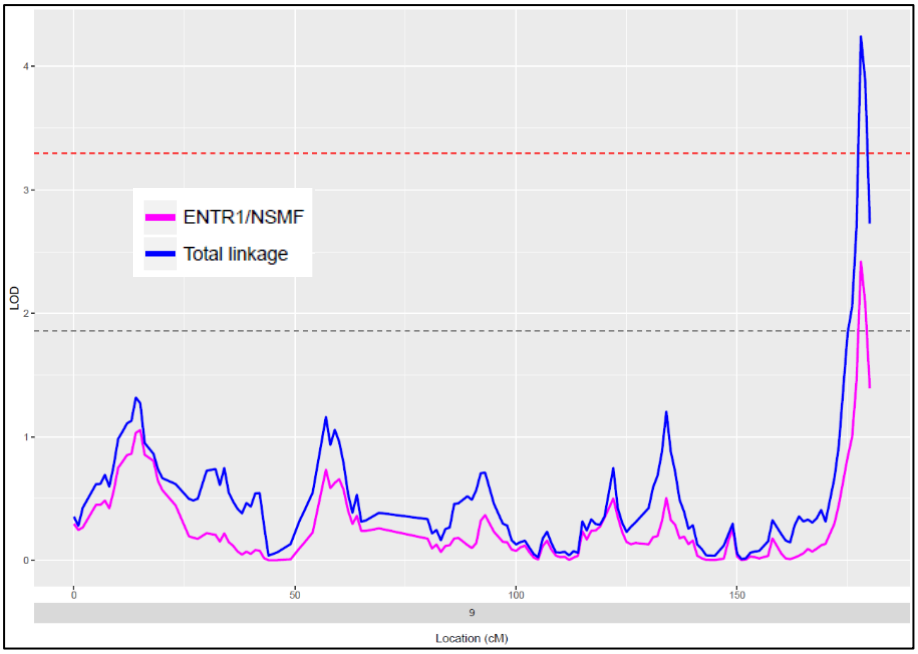


Figure 5-2 Linkage conditional on measured genotype of *ENTR1* or *NSMF* variants. Location (cM) = centimorgan location along each chromosome, **LOD** = logarithm of the odds, with the red dashed line indicating a significance level of 3.3 and the grey dashed line, a significance level of 1.86

5.3.4.2 Chromosome 15q11.2-13.2 linkage interval

The major contribution to the chromosome 15 peak was predicted to be from family 95002. Table 5-2 shows a reduction by a LOD score of 2 when this family was excluded from the linkage analysis. Figure 5-1f displays the whole linkage peak reduction, also showing a minor contribution to the linkage peak from families 93001 and GTAS54. The *HERC2* variant (novel variant located at chr15: 28474713) was identified in five members of family 95002 (Table 5-3) and was the only variant which fulfilled the prioritisation criteria. This synonymous SNP had the most significant MGA p-value of all the variants which were found predominantly in a single family (Appendix E) and had not been observed in the gnomAD database. This variant was associated with a 10mmHg increase in IOP, the largest increase of all the variants identified contributing to the linkage peaks. The variance explained by this SNP, at 4.6%, was the second highest in Table 5-3. When the linkage analysis was conducted, conditional on the measured genotype of the *HERC2* variant, the LOD score of the peak was reduced by 2, the greatest amount of all the variants assessed (Figure 5-3). In fact, this LOD reduction was almost identical to the linkage analysis when family 95002 was excluded (Figure 5-4).

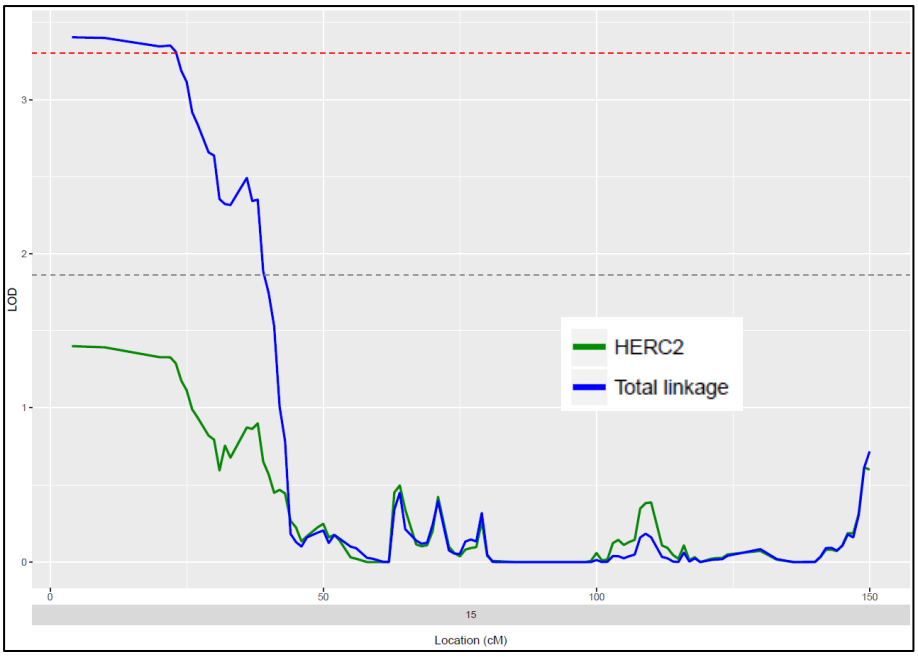


Figure 5-3 Linkage conditional on measured genotype of the *HERC2* variant. Location (cM) = centimorgan location along each chromosome, **LOD** = logarithm of the odds, with the red dashed line indicating a significance level of 3.3 and the grey dashed line, a significance level of 1.86

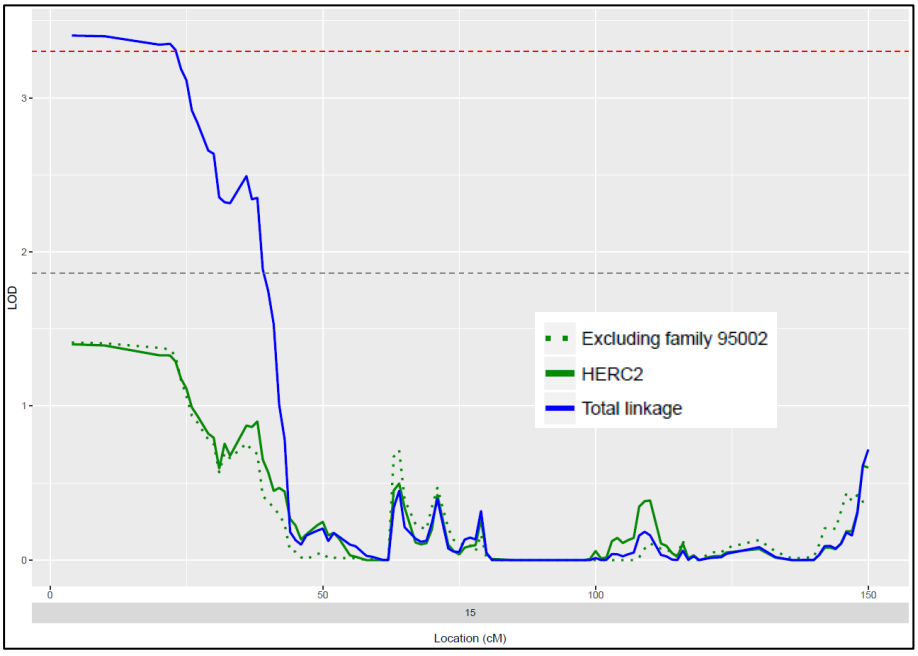


Figure 5-4 Linkage conditional on measured genotype of the *HERC2* variant compared to the exclusion of family 95002. Location (cM) = centimorgan location along each chromosome, **LOD** = logarithm of the odds, with the red dashed line indicating a significance level of 3.3 and the grey dashed line, a significance level of 1.86

The *HERC2* variant was not predicted to affect mRNA splicing and was not at an evolutionarily conserved site (Table 5-4). *HERC2* was expressed in the relevant ocular tissues in both the OTD and the RNAseq dataset (Table 5-5). The expression levels in both tissues of our in-house dataset were the second highest of all the assessed genes. The IPA analysis identified a protein interaction between *HERC2* and *TP53* [252], which was one of the POAG genes from the Janssen et al. (2013) dataset [13].

When the regions in and around *HERC2* was assessed in the MacGregor et al. (2018) [248], meta-analyses dataset the lead SNP reached suggestive significance ($p < 10^{-5}$) in the IOP GWAS. Figure 5-5 shows a LocusZoom plot of *HERC2* and surrounding genes, with the lead SNP reaching a significance of $p = 2.56 \times 10^{-6}$. This SNP was associated with a slight increase in IOP (beta-value = 0.07mmHg) and is located in an intronic region of *HERC2*. This SNP did not reach suggestive significance in the glaucoma GWAS, at $p = 5.44 \times 10^{-4}$.

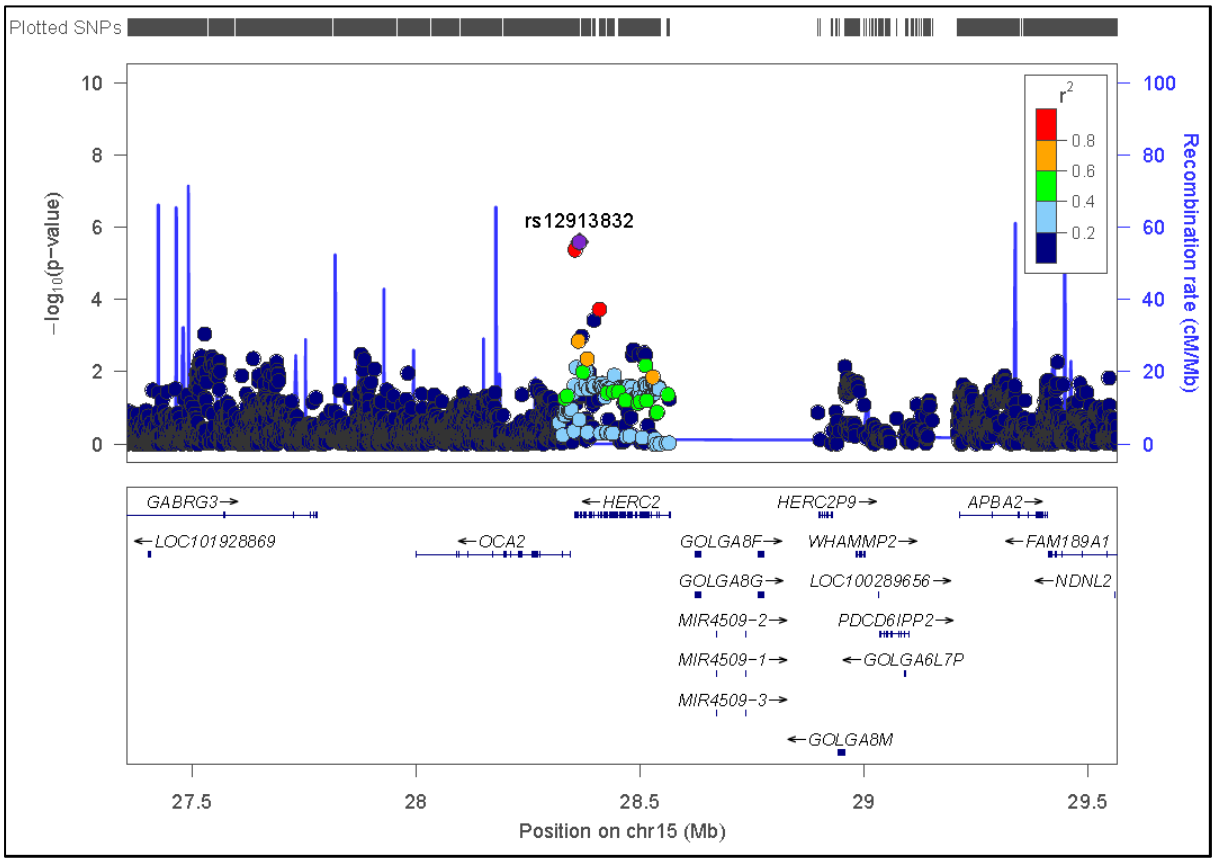


Figure 5-5 LocusZoom plot of part of the chromosome 15 linkage interval from the IOP GWAS, showing *HERC2*. The purple diamond represents the top statistically associated SNP. Pairwise correlations (r^2) between the top SNP and neighbouring SNPs are illustrated in different colours. Blue spikes show estimated recombination rates. cM = centimorgan, Mb = megabase

5.3.4.3 Chromosome 7p11.2-q11.23 linkage interval

The chromosome 7 linkage peak was the only peak to show involvement of only one family; 98002, the largest of the five families (Table 5-2 and Figure 5-1d). A SNP and indel (novel variants located at chr7: 71249529 and chr7: 71249770 respectively), only 250bp apart, in the 3' UTR region of *CALN1* were the only variants to fulfil the prioritisation criteria (Table 5-3). Both variants were extremely rare in the general population; the SNP had not been observed in gnomAD and the indel was seen with a frequency of less than 0.001. These variants are carried by 14 members of family 98002, 13 of whom had IOP phenotype data and contributed to the MGA results. These variants were associated with over 4mmHg increase in IOP and nearly 3% of the trait's variance. Figure 5-6 shows the effect of including the measured genotype of these variants on the chromosome 7 linkage peak. Most of the linkage peak was accounted for by these variants.

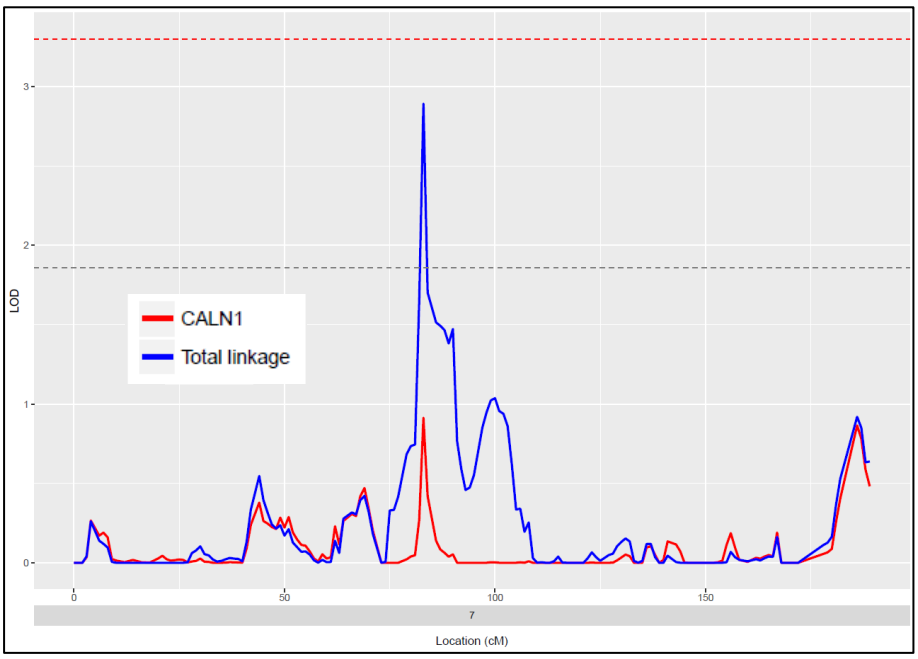


Figure 5-6 Linkage conditional on the measured genotype of the *CALN1* variants in family 98002. Location (cM) = centimorgan location along each chromosome, **LOD** = logarithm of the odds, with the red dashed line indicating a significance level of 3.3 and the grey dashed line, a significance level of 1.86

The *CALN1* variants were both predicted to affect mRNA splicing, and are both at evolutionarily conserved sites (Table 5-4). *CALN1* demonstrated a low level of expression in relevant ocular tissues in both datasets, especially in our RNAseq dataset, where the expression was minimal (Table 5-5). No interactions between this gene and the IOP and POAG genes

tested with IPA were identified and it did not reach significance in the published IOP and glaucoma meta-analyses.

5.3.4.4 Chromosome 2q23.1-31.1 linkage interval

The chromosome 2 linkage interval had five variants which fulfilled all the of the prioritisation criteria (Table 5-3). Three families contributed to this peak (Table 5-2 and Figure 5-1) and at least one variant in each of these families was identified which reduced the LOD score by a minimum of 1 when linkage was conducted conditional on their measured genotype. Figure 5-7 shows the effects of including the measured genotypes of these prioritised variants in the linkage model. A single line represents both of the *RIF1* variants (Figure 5-7b) as well as the *TTC21B* and *DLX1* SNPs (Figure 5-7c). Conducting linkage conditional on the measured genotypes of the variants from the three families together resulted in a complete reduction of the chromosome 2 linkage peak (Figure 5-7d). Each of these variants is discussed below under family sub-headings

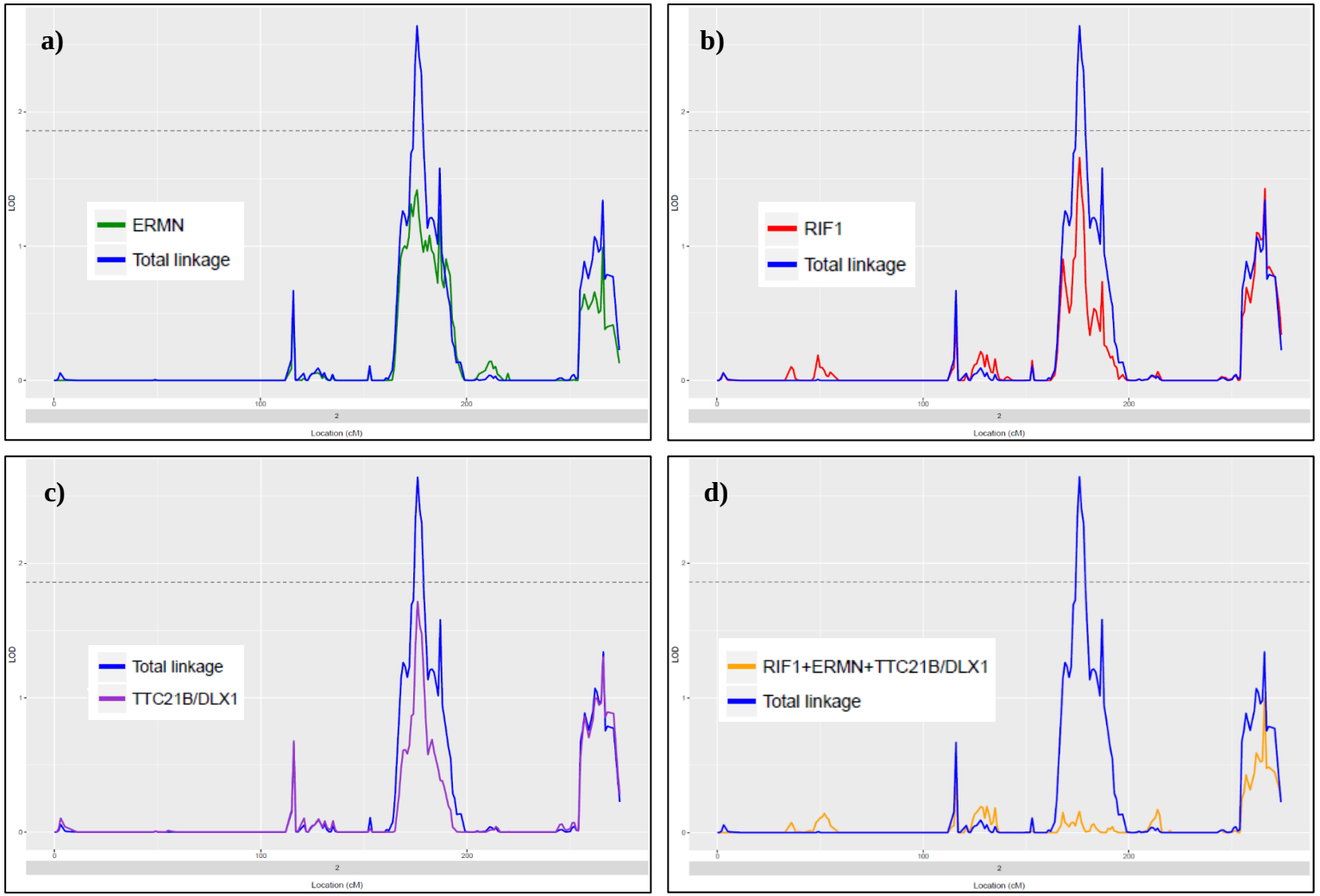


Figure 5-7 Linkage conditional on measured genotype of a) *ERMN*, b) *RIF1*, c) *TTC21B* or *DLX1* variants and d) combined *RIF1*, *ERMN* and *TTC21B* or *DLX1* variants Location (cM) = centimorgan location along each chromosome, **LOD** = logarithm of the odds, with the grey dashed line indicating a significance level of 1.86

Family 95002

The *ERMN* variant (rs rs749173309) was identified in 6 members of family 95002 (Table 5-3). Although this variant was found in the fewest number of individuals in the chromosome 2 linkage interval, it was associated with the greatest effect size, with an average 8mmHg increase in IOP per allele. This was also the rarest of the prioritised chromosome 2 variants, observed at a frequency of less than 0.0003 in the gnomAD NFE database. This variant also had the greatest reduction in maximum LOD of the chromosome 2 variants when linkage was conducted conditional on their measured genotype (Figure 5-2a). The *ERMN* variant was at an evolutionarily conserved site, with a GERP score of 3.3, and was also predicted to affect mRNA splicing by breaking an exonic splicing enhancer (ESE) site (Table 5-4). *ERMN* was not found in the OTD and it only showed minimal expression in the ciliary body and trabecular meshwork of our in-house RNAseq dataset (Table 5-5). This gene was not found to interact with any of the published genes assessed in the IPA analysis and did not reach significance in the published IOP and glaucoma meta-analyses.

Family 98002

Two variants in *RIF1*, less than 15kb apart, were identified in 34 members of family 98002 (two homozygous carriers resulted in a MAC of 36, Table 5-3). One of these variants is a synonymous SNP (rs116443402) and the other a deletion in the 3' UTR (rs372957656). These variants were more common than other prioritised variants in the general population, observed at a frequency of 0.02 in the gnomAD database. The *RIF1* variants are associated with a 1.7mmHg increase in IOP, a much smaller effect size than the *ERMN* variant. They also account for less trait variance (Table 5-3). The reduction in the linkage peak, conditional on the measured genotype of the *RIF1* variants (Figure 5-7b) is also not as great as that for the *ERMN* variant.

Two additional variants, in *NEB* (rs115631125) and *FMNL2* (rs138926754), were both found in the same individuals as the *RIF1* variants (Table 5-6). Although the MGA p-values of the variants in these genes were greater than 0.05, we hypothesised that there may be additional variants within the linkage region that influence IOP, common to the same set of family members.

Table 5-6 Additional variants identified in the same individuals as the *RIF1* variants in Family 98002
Variants in bold are additional to those already shown in Table 5-3 (not bold)

										MGA				Family specific MAC			
Chr	**bp position (hg19)**	**Variant ID**	**Gene**	**Gene region**	**Ref**	**Alt**	**CADD**	**gnomAD genome NFE**	**AF**	**p-value**	**beta-value**	**variance explained (%)**	**MAC**	**93001**	**95002**	**98002**	**GTAS54**
2	152,317,669	rs116443402	*RIF1*	exonic	G	A	0.3	0.020	0.072	0.0215	1.71	1.74	34			36	
2	152,331,995	rs372957656	*RIF1*	3' UTR	GTTG	-	1.6	0.020									
2	**152,417,767**	**rs115631125**	***NEB***	**exonic**	**C**	**A**	**< 0.1**	**0.004**	**0.042**	**0.2047**	**1.24**	**0.70**	**20**			**21**	
2	**153,378,462**	**rs138926754**	***FMNL2***	**exonic**	**T**	**C**	**10.9**	**0.029**	**0.058**	**0.1857**	**0.34**	**0.45**	**28**	**1**		**25**	**3**

Chr = chromosome, **bp** = base pair
3' UTR - 3' untranslated region, **5' UTR -** 5' untranslated region, **exonic** = protein coding region (synonymous)
Ref = reference allele, **Alt** = alternate allele
CADD = combined annotation dependent depletion phred scores (version 1.4)
gnomAD genome NFE = variant allele frequency in non-Finnish Europeans in the gnomAD genome database.
AF = frequency of the variant allele in the 5 families
MGA = measured genotype association analysis
beta-value = effect size per allele; pressure inside the eye measured in mmHg,
MAC = minor allele copies as used for MGA analysis.
Family specific MAC = minor allele copies in each family. May differ from MAC used for MGA if IOP information was not available for particular individuals

Figure 5-8 shows comparisons of linkage conditional on the measured genotypes of the *RIF1* variants with *NEB* (Figure 5-8a) and *FMNL2* (Figure 5-8b) variants. Both of these variants have very similar effects on the chromosome 2 peak as the *RIF1* variants, reducing the maximum LOD score by 1 (*NEB*) and 0.9 (*FMNL2*).

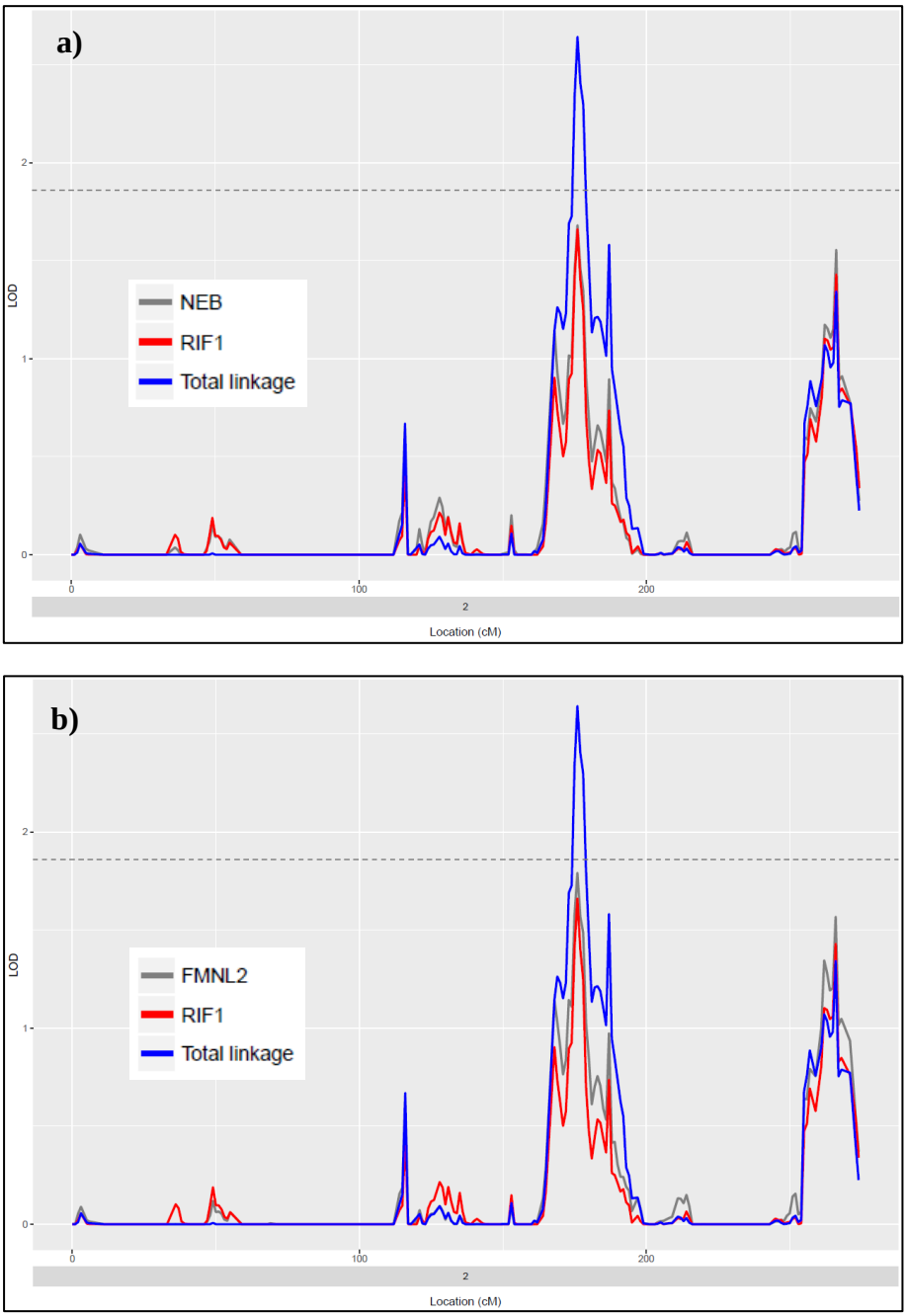


Figure 5-8 Linkage conditional on the measured genotype of the *RIF1* variants versus the NEB variant (a) and *FMNL2* variant (b). Location (cM) = centimorgan location along each chromosome, **LOD** = logarithm of the odds, with the grey dashed line indicating a significance level of 1.86

Of the members of Family 98002 shown in Table 5-6 carrying variants in the *RIF1*, *NEB* and *FMNL2* genes, 19 were carriers common for all the variants across the 3 genes. When linkage was conducted using the measured genotypes of only the 19 members of Family 98002 who carry the variants in these three genes, the maximum LOD score was reduced even further than for the *RIF1* variants alone (Figure 5-9). This suggests that these 19 family members were driving the peak, and the additional family members who were carriers of the *RIF1* variants (Table 5-6) were not important to this linkage region.

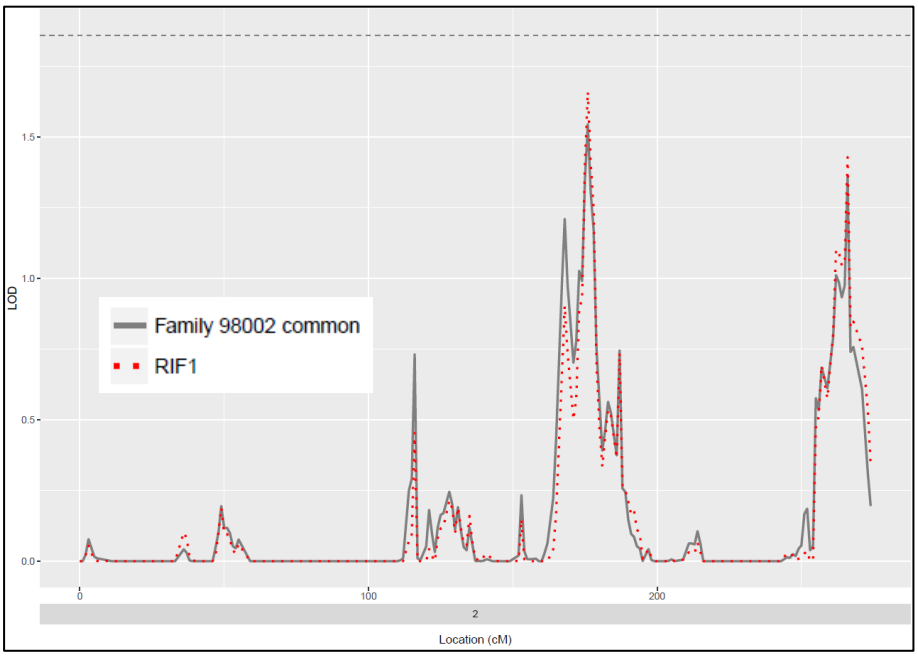


Figure 5-9 Linkage conditional on the measured genotype of the *RIF1* variants versus the measured genotype of the *RIF1*, *NEB* and *FMNL2* variants common to 19 members of Family 98002. Location (cM) = centimorgan location along each chromosome, **LOD** = logarithm of the odds, with the grey dashed line indicating a significance level of 1.86

Of the four variants identified in family 98002 in the chromosome 2 linkage interval, three were predicted to possibly affect mRNA splicing (Table 5-4), with only the *RIF1* exonic SNP with no predicted effect. *RIF1*, *NEB* and *FMNL2* are all expressed in the ciliary body and trabecular meshwork tissues of the eye, with *NEB* demonstrating a much greater expression in the trabecular meshwork than the ciliary body in both the OTD and our in-house RNAseq data (Table 5-5). The expression of *NEB* in the trabecular meshwork is the greatest of all the genes in this study which were assessed in the OTD, with a plier score of 129. The IPA gene interaction analysis revealed an interaction at the protein level between *FMNL2* and *CDH1*

[253]. *CDH1* is a member of the Type 1 classical cadherin group, involved with cell adhesion (https://www.genenames.org/).

When variants in the region encompassing *RIF1*, *NEB* and *FMNL2* were assessed for significance in the IOP and glaucoma meta-analyses [248], genome-wide significance, at $p < 5x10^{-8}$, was reached within *FMNL2*. The lead SNP in *FMNL2* (rs1579050) was highly significant at $p = 1.39x10^{-25}$. Figure 5-10 shows a plot of the chromosome 2 linkage interval encompassing *RIF1*, *NEB* and *FMNL2*. The lead SNP is located in the first intron of *FMNL2* and the risk allele was associated with an increase in IOP (beta-value = 0.13mmHg). In the glaucoma GWAS, this locus did not quite reach genome-wide significance, with rs1579050 reaching $p = 2.90x10^{-7}$. The glaucoma odds ratio for the risk allele was 1.08 (95% CI = 1.05, 1.11).

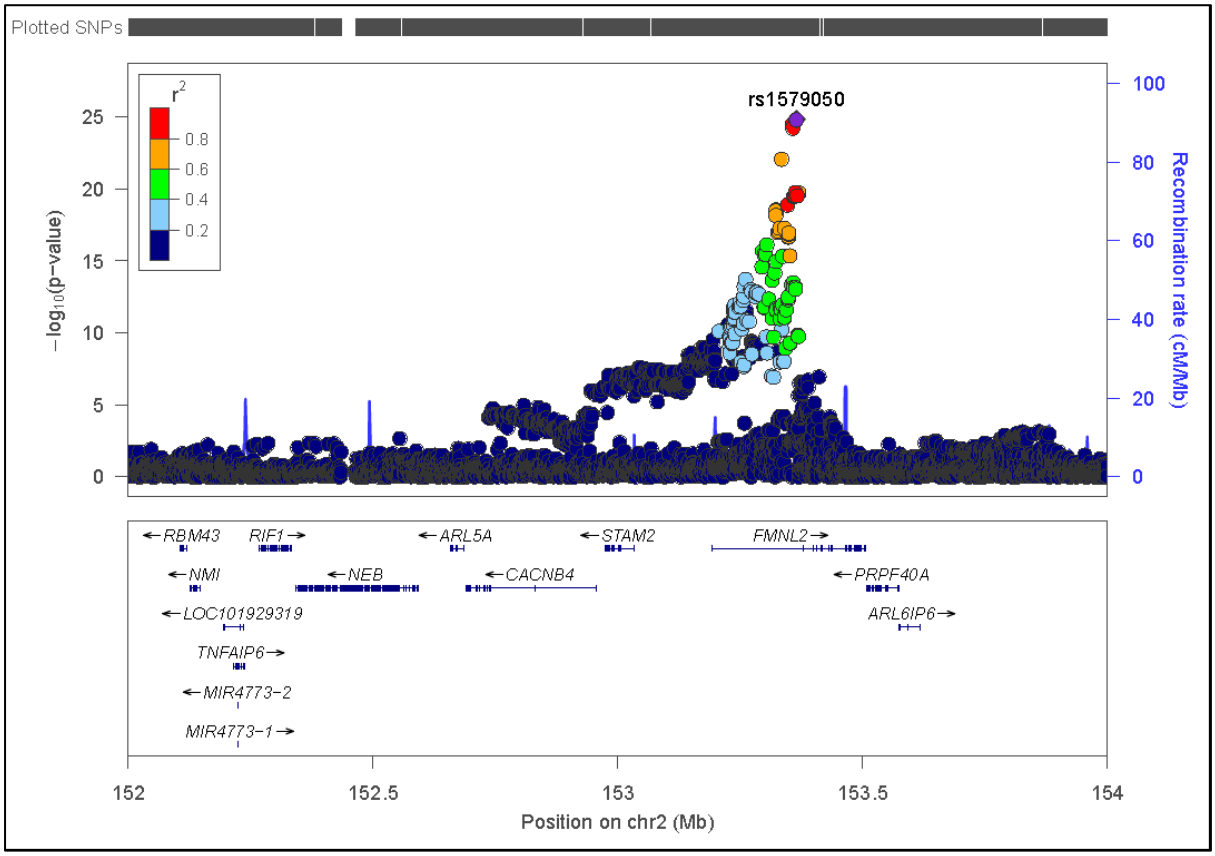


Figure 5-10 LocusZoom plot of part of the chromosome 2 linkage interval from the IOP GWAS, encompassing *RIF1*, *NEB* and *FNML2*. The purple diamond represents the top statistically associated SNP. Pairwise correlations (r^2) between the top SNP and neighbouring SNPs are illustrated in different colours. Blue spikes show estimated recombination rates. cM = centimorgan, Mb = megabase

Family GTAS54

A SNP each in *TTC21B* (rs572065013) and *DLX1* (rs186854626) were identified in the same nine GTAS54 family members (Table 5-3). These variants were not common in the gnomAD database, observed at a frequency of less than 0.005. These variants contribute to nearly 4% of the IOP trait variance and were associated with a 5mmHg increase in IOP in these family members. The *DLX1* variant was predicted to affect splicing and is at an evolutionarily conserved site, with a GERP score of 3.7, whereas the *TTC21B* variant was not predicted to affect splicing and is not evolutionarily conserved (Table 5-4). These two variants are over 6 million bp apart, and as these variants were found in the same family members, it is impossible to determine from the MGA and linkage results which variant to prioritise over the other. But it does suggest that there may be a large haplotype in this family which may contribute to IOPmed variance.

Another variant which was present only in Family GTAS54, with an MGA p-value less than 0.05 was in *PKP4* (rs148782148, Table 5-7). This variant was not prioritised with *TTC21B* and *DLX1*, as when linkage was conducted conditional on the measured genotype of the *PKP4* variant, the maximum LOD was reduced by 0.7, not quite reaching the threshold required for prioritisation of variants. This SNP is over 7 million bp away from *TTC21B* and was found in eight out of the nine GTAS54 family members who carry the prioritised variants in *TTC21B* and *DLX1*. This non-synonymous SNP had the highest CADD score of all the assessed variants in all of the linkage intervals, at 28.3. It was also at a highly evolutionarily conserved site, and with a GERP score of 5.5, was the most conserved of all the assessed variants (Table 5-4). This variant was also predicted to affect mRNA splicing with the insertion of a new exonic splicing silencer (ESS) site as well as breaking an exonic splicing enhancer (ESE) site (Table 5-4). The *PKP4* variant contributed to more of the IOPmed trait variance than the *TTC21B* and *DLX1* variants, at nearly 6%, and had a similar associated effect size. The expression of *PKP4* was the highest in the ciliary body of all the genes assessed, across all the linkage regions, in both the OTD (plier score of 158) and our in-house RNAseq data (184 counts per million) (Table 5-5). *TTC21B* and *DLX1* showed lower expression in the OTD and a much lower expression level in our RNAseq dataset, with *DLX1* with only minimal expression (< 1 count per million). In the IPA gene interaction analysis, a protein interaction between *PKP4* and *CDH1* was identified [254]. *CDH1* was the same gene identified as interacting with *FMNL2*, from family

98002 described above. None of these genes identified in Family GTAS54 reached significance in the published IOP and glaucoma meta-analyses.

Table 5-7 Additional variant identified in the same individuals as the *TTC21B* and *DLX1* variants in Family GTAS54
The variant in bold is additional to those already shown in Table 5-3 (not bold)

										MGA				Family specific MAC			
Chr	**bp position (hg19)**	**Variant ID**	**Gene**	**Gene region**	**Ref**	**Alt**	**CADD**	**gnomAD genome NFE**	**AF**	**p-value**	**beta-value**	**variance explained (%)**	**MAC**	**93001**	**95002**	**98002**	**GTAS54**
2	**159,536,990**	**rs148782148**	***PKP4***	**exonic p(Asp>Val)**	**A**	**T**	**28.3**	**0.005**	**0.016**	**0.0013**	**4.6**	**5.76**	**8**				**8**
2	166,810,232	rs572065013	*TTC21B*	5' UTR	G	A	6.4	0.004	0.018	0.0079	5.02	3.96	9				9
2	172,953,525	rs186854626	*DLX1*	3' UTR	C	T	9.9	0.002									

Chr = chromosome, **bp** = base pair
3' UTR - 3' untranslated region, **5' UTR** - 5' untranslated region, **exonic** = protein coding region
Ref = reference allele, **Alt** = alternate allele
CADD = combined annotation dependent depletion phred scores (version 1.4)
gnomAD genome NFE = variant allele frequency in non-Finnish Europeans in the gnomAD genome database.
AF = frequency of the variant allele in the 5 families
MGA = measured genotype association analysis
beta-value = effect size per allele; pressure inside the eye measured in mmHg,
MAC = minor allele copies as used for MGA analysis.
Family specific MAC = minor allele copies in each family. May differ from MAC used for MGA if IOP information was not available for particular individuals

5.3.4.5 Chromosome 3p13-12.1 linkage interval

Family GTAS54 was predicted to contribute most to the chromosome 3 linkage peak, followed by family 93001 and 95002 (Table 5-2). Two variants, in *PDZRN* (rs61732653) and *FAM86DP* (rs184945375), fulfilled the criteria for prioritisation (Table 5-3). These SNPs were found in the same nine GTAS54 family members, with an additional family member carrying the *FAM86DP* variant, and a family 98002 individual carrying the *PDZRN* variant. These two variants are located 2 million base pairs apart, with no other variants found predominantly in the same GTAS54 family members within the chromosome 3 linkage interval. Both variants were associated with a similar increase in IOP, over 4mmHg per allele, with the *PDZRN* variant explaining slightly more of the variance. When linkage was conducted conditional on the measured genotypes of these variants, the *PDZRN* variant demonstrated a larger reduction in the maximum LOD (Figure 5-11). Excluding the family 98002 individual from the analysis had a negligible effect on the peak (linkage plot not shown).

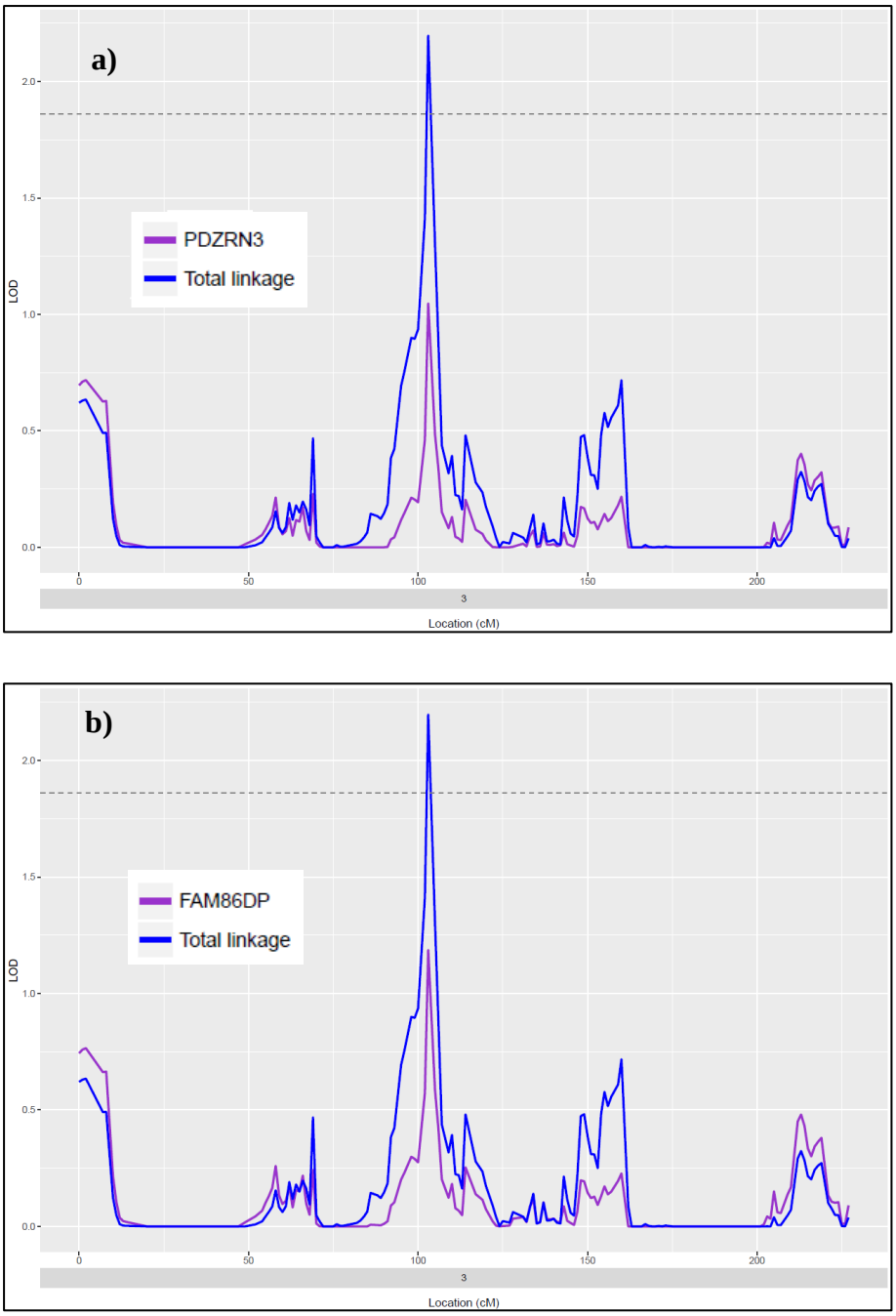


Figure 5-11 Linkage conditional on the measured genotype of the a) *PDZRN* variant and b) *FAM86DP* variant. Location (cM) = centimorgan location along each chromosome, **LOD** = logarithm of the odds, with the grey dashed line indicating a significance level of 1.86

When assessing other features of these two variants, both were predicted to affect mRNA splicing and both sites were not evolutionarily conserved (Table 5-4). *PDZRN* was more highly expressed than *FAM86DP* in both the ciliary body and trabecular meshwork ocular tissues in both the OTD and the RNAseq dataset (Table 5-5). Neither of these genes demonstrated any interactions with the other IOP and POAG genes assessed in the gene interaction analysis and

neither of them demonstrated any significant associations in the IOP and glaucoma meta-analyses.

5.3.4.6 Chromosome 6q24.2-25.1 linkage interval

The chromosome 6 linkage peak did not have a major contribution by a single family (Table 5-2). Family 95002 was shown to contribute most to this peak, albeit a small contribution. When this linkage interval was investigated to prioritise variants, only three variants were found predominantly in family 95002. Two of these variants had MGA p-value ≤ 0.05 (Appendix E), and six of the eight family members were common to the two variants. However, when tested for linkage conditional on their measured genotype, the best of these two variants, rs143818607 in *PLEKHG1*, only demonstrated a maximum LOD reduction of 0.4. No variants fulfilled all of the prioritisation criteria for this linkage region.

5.4 Discussion

Six QTLs for medication adjusted IOP were identified in five families as described in the previous chapter. The purpose of the analyses described in this chapter was to identify the variants driving the peaks and propose genes which may be involved in IOP regulation in the families of this study. Variants were prioritised within the linkage peaks by firstly, restricting the analysis to variants present in the families making the major contribution to each peak with the most significant MGA results. Variants meeting these criteria and reducing the LOD score by at least 1 LOD unit in conditional linkage analysis were considered prioritised. Additional variants, found in the majority of the same family members as prioritised variants, were also assessed to determine whether there could be alternative risk variants within a haplotype. In total, 15 variants in 13 genes across all of the QTLs, had a major influence on the magnitude of the peak linkage when adjusted for in the linkage model. No variants affecting peak linkage were identified in the chromosome 6 linkage region whereas eight variants affecting linkage were found in the chromosome 2 linkage region.

One of the most important findings in this study is the chromosome 15 linkage region. This was one of the two peaks which reached full statistical significance, with a maximum LOD score of 3.40 after correction for the ascertainment of POAG families and adjustment for the non-normal distribution of the IOPmed trait (detailed in Chapter 4). This is the only linkage region in our study which overlaps a previously published linkage region. The peak identified in this study completely overlaps the GLC1I region, identified by ordered subset analysis in 15 families with early onset POAG [177]. The fact that 15 families contributed to that linkage peak suggests a more common variant was involved or possibly several rare variants across multiple families. The *HERC2* variant identified in our study is so rare that it has not been observed in the gnomAD database in any population. The linkage region in our study was driven predominantly by one family (Table 5-2 and Figure 5-1f), which contributed a LOD score of 2 to this peak. A variant in *HERC2*, identified in five members of family 95002, accounted for the whole contribution to the peak (Figures 5-3 and 5-4). As a potentially private variant in family 95002, it is highly unlikely to be the same functional variant driving the linkage peak at the GLC1I locus. However, there may be other rare variants in *HERC2* in the 15 families included in the GLC1I locus investigation, and further screening of those families is necessary.

The MGA results in our study demonstrated association of the *HERC2* variant with a large effect on the average IOP in these five family members and which accounted for over 4.5% of the trait variance (Table 5-3). In the meta-analysis conducted by MacGregor et al. (2018) [248], the *HERC2* locus reached suggestive significance for association with IOP ($p = 2.56x10^{-6}$ for the lead SNP, rs12913832). Although that study identified the *HERC2* locus as associated with increased IOP, it does not confirm this gene as important in IOP regulation. Our study has identified a coding variant within the *HERC2* gene itself and is a step towards confirming the possible involvement of *HERC2* with IOP regulation. Comparing the common lead SNP in the meta-analysis (MAF=0.56) and its associated effect of increasing IOP by 0.07mmHg, with the rare variant identified in our study, associated with over 10mmHg increase in IOP, demonstrates the benefit of using families to look for rare variants with large effect sizes.

HERC2 was highly expressed in both the ciliary body and trabecular meshwork of our RNAseq database. This gene is involved in DNA repair [255] and some variants in *HERC2* are involved with eye, skin and hair pigmentation [256-258] and others are associated with intellectual disability [259, 260]. In the gene interaction analysis, a protein interaction between *HERC2* and *TP53* was identified. *TP53* was previously associated with POAG [261-264], although results have been conflicting, with other researchers not finding an association [265, 266]. Tumour protein 53 (p53) is an important transcription factor which regulates the cellular response to stress [267] and the previous studies which found an association with POAG suggested that the role of p53 in apoptosis may be relevant to the retinal ganglion cell death which is characteristic of POAG pathogenesis [261-264]. The HERC2 protein has been proposed to regulate p53 signalling by binding to the protein and activating p53 oligomerisation, an integral part of regulating the transcription of target genes [252, 268]. However, a possible role that *HERC2* and *TP53* may play in apoptosis together is still to be elucidated. Although *TP53* has been associated with POAG, no role for this gene in IOP regulation has been proposed. *HERC2* is a novel gene for IOP identified in our current study, and very recently the locus has been (independently) identified by GWAS [83], however, its mechanism of action in IOP regulation is unknown.

Figure 5-13 shows the POAG disease status of family 95002 and identifies the *HERC2* variant carriers. Of these five individuals who carry the *HERC2* variant, four of them were diagnosed with POAG when clinically assessed. Only one of the five variant carriers did not have a POAG diagnosis at their last clinical assessment. However, this person (9500217) had high IOP

(27mmHg, well above the ascertainment correction population average of 15mmHg), and may well have developed POAG since the study concluded as they were only 55yrs old at their last clinical assessment. Although this is not enough evidence to state a link between this variant or gene with POAG, it is a promising finding and worthy of further investigation.

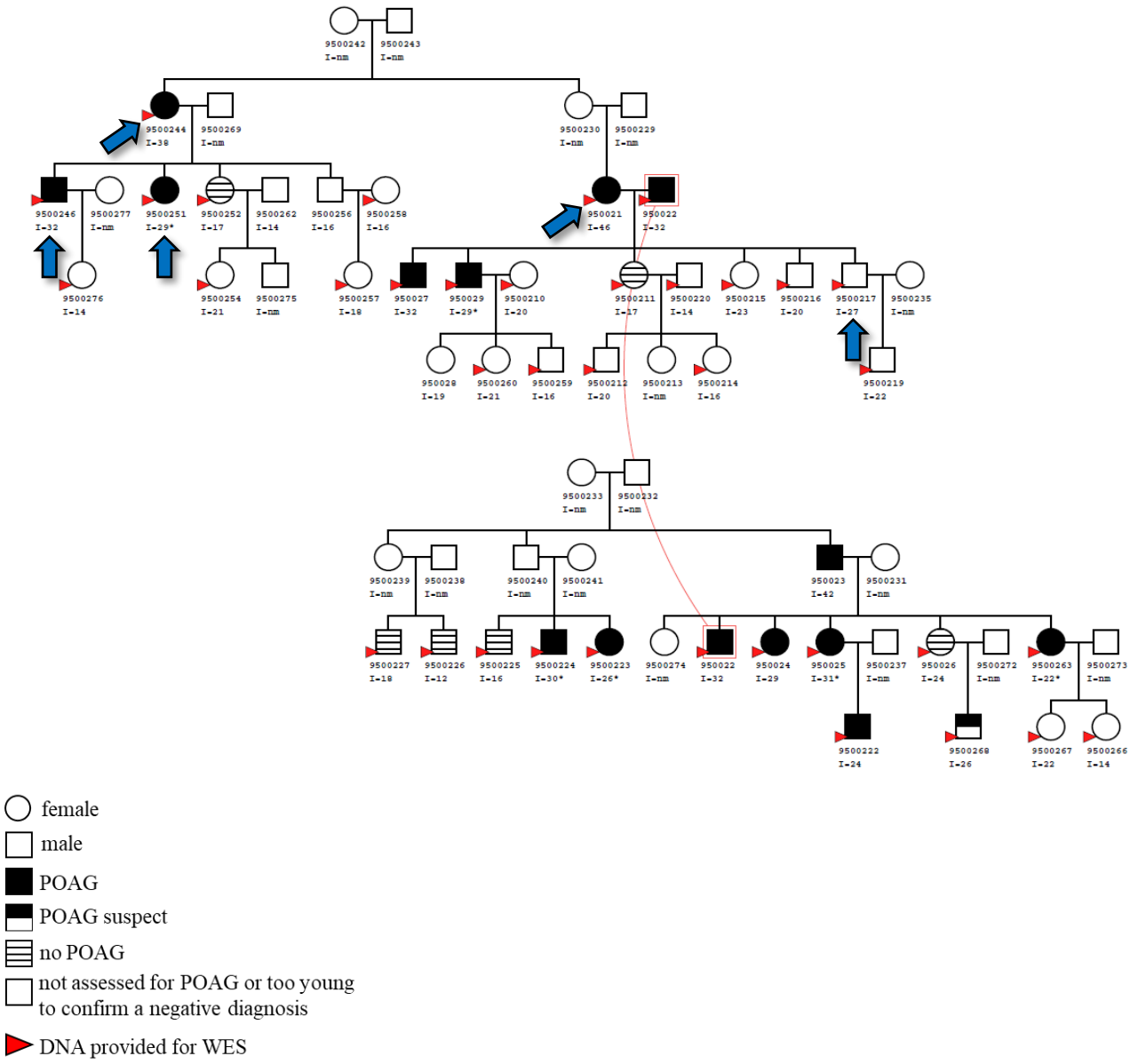


Figure 5-12 Family 95002 indicating POAG disease status and carriers of the novel *HERC2* variant, at 15:28474713 (chr:pos hg19) Blue arrows indicate variant carriers, red curved line indicates the same individual. Identification number = the first string of numbers **I** = IOP measurement (mmHg), ***** = medication adjusted IOP

Although the *HERC2* variant identified here is exonic, associated with a large increase in IOP and the gene is expressed in relevant ocular tissues, it may not be a functional variant itself. The CADD score of 10 does not predict this variant to be highly deleterious, but functional research on this gene would be required to determine which variants affect gene function. The gnomAD database proposes *HERC2* as a highly constrained gene, intolerant of missense and loss of function variants. Other, more deleterious, variants may be present on the same haplotype in the same five family members as the identified variant and further sequencing of other *HERC2* regions in these five individuals may lead to other variants of interest.

As with the chromosome 15 linkage region, variants with the potential to affect IOP regulation have also been identified within the chromosome 2 QTL. However, where the chromosome 15 peak was obviously driven by one family, the chromosome 2 peak is more complex and influenced by multiple families. Excluding individual families from the linkage analysis suggested Families 95002, 98002 and GTAS54 as contributing to this region (Table 5-2). The *ERMN* variant, carried by 6 members of family 95002, had the largest effect on reducing the maximum LOD (Figure 5-7a). Although this 3' UTR variant demonstrated association with increased IOP in these family members in the MGA results (Table 5-3) the CADD score did not predict it to be deleterious. *ERMN* was not highly expressed in the relevant ocular tissues (Table 5-5) and a link with IOP is difficult to determine. The true functional variant behind this family's effect on the linkage peak may be on the same haplotype as this *ERMN* variant, but at this stage, cannot be identified.

The family 98002 contribution to the chromosome 2 linkage peak was identified as a presumed haplotype encompassing at least the *RIF1*, *NEB* and *FMNL2* genes. The *RIF1* variants were prioritised based on their MGA results and reduction in maximum LOD when included in the linkage model. Additional variants were identified in *NEB* and *FMNL2* which were carried by many of the same individuals as the *RIF1* variants. Nineteen members of family 98002 carried variants in all three of these genes and the measured genotype of only these individuals reduced the linkage peak by a greater amount than the 34 *RIF1* carriers alone (Figure 5-9). Further analysis will need to be undertaken to confirm whether this region on chromosome 2 is indeed the same haplotype in these 19 individuals. Currently, the region identified here encompasses over 1 million base pairs and includes the genes; *ARL5A*, *CACNB4* and *STAM2*, which did not have any prioritised variants in this study. The presumed haplotype could well extend beyond the genes identified here.

RIF1 was the original gene with the identified variants from Family 98002 in the chromosome 2 linkage interval. However, although this gene is expressed in the eye (Table 5-5), it has not been associated with any eye diseases. *NEB* is expressed in the IOP relevant tissues of the eye and particularly in the trabecular meshwork more so than the ciliary body (Table 5-5). This gene is a component of muscle tissue and is a causative gene of nemaline myopathy, a neuromuscular disorder [269]. Interestingly, an association was found between a non-synonymous *NEB* variant and primary angle closure glaucoma (PACG) in a breed of dog susceptible to this form of glaucoma [270]. PACG is also associated with an increase in IOP, and the mechanism of involvement with glaucoma in the dog study is believed to be due to an effect on the ciliary muscle of the eye which is necessary to retain normal structure of the anterior chamber of the eye including regulating the outflow of aqueous humour. However, there is currently no evidence that *NEB* is involved with IOP regulation in humans.

FMNL2 is the most likely candidate gene to affect IOP out of the three identified in Family 98002. Although the particular variant identified here was not prioritised based on the MGA results and it barely contributed to the variance of IOP in the families of this study (Table 5-6), *FMNL2* has already been associated with IOP [111, 171] and POAG [111] (also see Appendix B). The two published variants were identified from large GWAS and although not proposed to be the functional variants themselves, the authors proposed *FMNL2* as a potentially important IOP and POAG gene [111]. After association with this gene was found, functional studies were conducted to determine whether *FMNL2* affects trabecular meshwork function which could potentially affect the regulation of aqueous humour outflow [111]. That study suggested that knockdown of the *FMNL2* gene indirectly affected the assembly of actin fibres, via cell-adhesion based means, which induced morphological changes in human trabecular meshwork cells. They proposed that the effect on the contractile properties of trabecular meshwork cells could affect aqueous humour outflow, with a resultant increase in IOP as the mechanism of POAG pathogenesis. Strengthening this link between *FMNL2* and cell adhesion mechanisms of action were the results of the gene interactions analysis conducted in our study. The IPA analysis identified a direct protein-protein interaction between FMNL2 and CDH1, also known as E-cadherin. These two proteins together are an important component of the cell-cell adhesion complex [253]. Other studies have reported the importance of cadherins in cell-cell adhesion in trabecular meshwork tissues, the possible dysfunction of which is involved in IOP dysregulation and glaucoma, although these studies focused on K-cadherin [271] and N- and OB-cadherins [272]. The *CDH1* gene was found to be upregulated in retinal ganglion cells

in a rat model with ocular hypertension [273] and was proposed as an important gene in the pathogenesis of POAG [13].

FMNL2 is also a promising candidate gene for IOP regulation and POAG pathogenesis when the significance of this locus was assessed in the meta-analysis conducted by MacGregor et al. (2018) [248]. That study identified the *FMNL2* locus as highly significantly associated with IOP (Figure 5-19) as well as being associated with glaucoma. The evidence surrounding *FMNL2* as a gene involved in IOP regulation and glaucoma is encouraging and supports the proposal of this gene as the most likely candidate in the chromosome 2 linkage region for Family 98002. Although the variant identified from Family 98002 in this current study is unlikely to be the functional variant itself, it is only 70kb away from the two variants discussed above [111, 171]. Unfortunately, poor sequencing quality around this region in our study resulted in several variants being excluded from the MGA analysis, and resequencing this region may identify other variants with potentially functional consequences.

Along with Families 95002 and 98002, GTAS54 also contributed to the chromosome 2 QTL and potentially supplies a novel and interesting candidate gene. Of the three variants shown in Table 5-7, the *PKP4* variant is prioritised over the *TTC21B* and *DLX1* variants. This is the only non-synonymous protein coding variant, predicted to be highly deleterious, with the highest CADD score of all the variants identified, in any of the linkage regions. It is predicted to potentially affect splicing of the mRNA and is the most evolutionarily conserved of the three variants. *PKP4* is the most highly expressed gene, of assessed genes in this study, in the ciliary body in both the OTD and our in-house RNAseq database, far exceeding the expression levels of the other genes in this tissue (Table 5-5), and it is also expressed in the trabecular meshwork, with its expression in the OTD the second highest of the genes assessed.

PKP4 is a member of the catenin protein family, which are involved with intercellular adhesion through the stabilisation of cadherins, the molecules which bind cells together [274, 275]. K-cadherin is highly expressed in normal trabecular meshwork tissues and is regulated by the wnt/β-catenin signalling pathway. K-cadherin contributes to cell-cell adhesion in the trabecular meshwork is believed to be important in the regulation of IOP. Blocking of the wnt/β-catenin pathway led to an increase in IOP in a mouse model of glaucoma [271]. Although a direct relationship between PKP4 and K-cadherin in the trabecular meshwork has not yet been established, PKP4 is known to interact with other types of cadherins and is believed to be important in cell-cell adhesion as well as other cellular functions [274]. The IPA gene

interactions analysis conducted in our study supports a cell adhesion effect model, with an interaction between PKP4 and CDH1 (E-cadherin) identified [254], which is the same protein identified as interacting with FMNL2 discussed above. As *PKP4* expression in the ciliary body is higher than even the trabecular meshwork, there may also be a role in this tissue, which produces the aqueous humour responsible for maintaining IOP in the eye. In a study based on a mouse knock-down model of p120 catenin, a member of the same protein family as PKP4 (the p120 catenin family), severe developmental defects were observed in the anterior segment of the eye, including ciliary body and trabecular meshwork tissues [276]. Further research on PKP4 and its role in cell adhesion and IOP regulation in both the ciliary body and trabecular meshwork is necessary. There are five isoforms of *PKP4* (NM_003628.5, NM_001005476.3, NM_001304969.2, NM_001304970.2 and NM_001304971.1) and the variant identified in this study only affects four of them (not NM_001304971.1). The summary RNAseq data available to our study did not provide the level of detail necessary to confirm which transcripts were present in the ocular tissues. It is necessary to identify which mRNA transcripts are expressed in the relevant tissues of the eye to determine whether the *PKP4* variant could potentially play a functional role in IOP regulation.

Figure 5-14 shows the POAG disease status of family GTAS54 and identifies the *PKP4* variant carriers. Of the eight people identified in family GTAS54 who carry this PKP4 variant, four had a confirmed POAG diagnosis, three were suspected to have POAG and one had no diagnosis at their last clinical examination. This person (1158644) had elevated IOP of 24mmHg (well above the ascertainment correction population mean of 15mmHg), but was not yet exhibiting optic nerve damage characteristic of glaucoma. At 49 years old at their last clinical examination, this person may have since developed POAG, but this is currently unknown. Although this is not enough evidence to state a link between this variant or gene with POAG, it is a promising finding and worthy of further investigation.

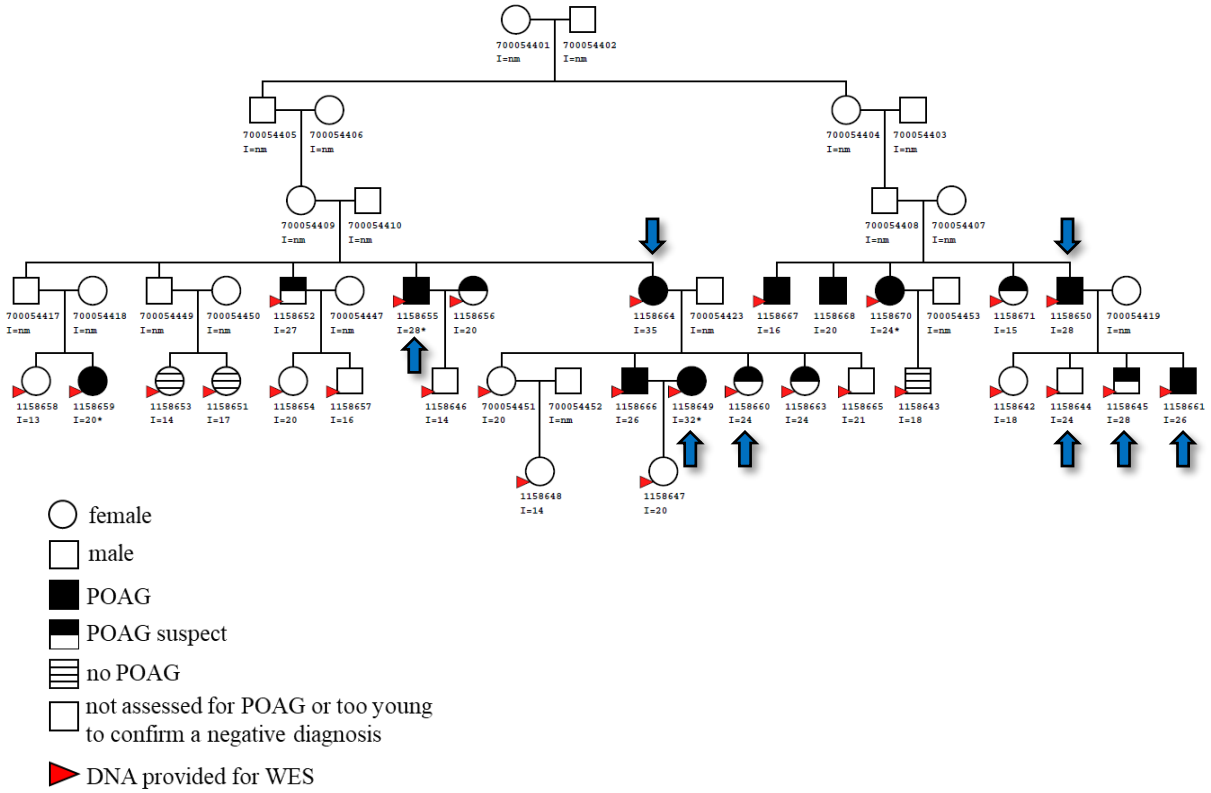


Figure 5-13 Family GTAS54 indicating POAG disease status and carriers of the PKP4 variant, rs148782148. Blue arrows indicate variant carriers. Identification number = the first string of numbers **I** = IOP measurement (mmHg), ***** = medication adjusted IOP

The chromosome 2 QTL is complex, with three families contributing to the observed linkage signals. This peak reached the suggestive significance threshold (LOD=1.86), which was used to select linkage regions identified in Chapter 4, due to the contribution of the three families together. Although the families' linkage signals did not reach suggestive significance individually, that does not mean the linkage peak is not worth further consideration. Lander and Kruglyak (1995), propose that linkage peaks are still worth investigating with a nominal p-value = 0.05 which is equivalent to a LOD score of 0.59, as long as it is understood that there will be false positive results obtained [232]. Using the reduction in LOD, when linkage was conducted conditional on the measured genotype of these variants, as a proxy for the LOD score of that family's contribution, allows each of these families' contributions to be assessed. At least one variant was identified in each family which reduced the LOD score by at least 1 LOD unit (Table 5-3). This linkage region may represent a true and coincidental finding of three families with IOP regulating genes within close proximity to each other. Conversely, this region may represent one or more false positive signals. Further research is necessary to confirm whether each of the three families do indeed provide a true linkage signal in this region. When each of the families were analysed separately, promising candidate genes were identified in two of the families. *FMNL2* in Family 98002 has already been associated with IOP and *PKP4*, in Family GTAS54, is a novel, but plausible candidate for IOP.

Of the other linkage regions assessed in this study, there is less evidence for the identification of candidate genes than for the chromosome 2 and 15 regions. The chromosome 9 linkage peak was the most significant, reaching a LOD score of 4.25. Family 93001 was predicted to contribute most to this peak and two variants identified in the same members of this family, over a million bp apart, were responsible for the largest reduction in the linkage peak (Figure 5-2). Thus there is still part of this family's contribution to the peak which is unaccounted for. Of the two genes in which variants were identified; *NSMF* and *ENTR1*, the variant in *NSMF* had the highest CADD score and this gene had greater expression in the ciliary body and trabecular meshwork in our RNAseq database (*NSMF* was not included in the OTD). *NSMF* has a well-recognised role in growth and migration of specific neuronal cells [277], but as yet, no involvement with the eye. The fact that two variants spanning a region over 1Mb, including over 50 genes, were identified in the same individuals suggests that these variants could be tagging the true functional variant which has not been sequenced in the WES of this study. No IOP or POAG genes have been identified in this region of chromosome 9 previously and further analysis is necessary.

The QTL on chromosome 7 was the only one driven by a single family (Table 5-2 and Figure 5-1d). Two variants in *CALN1*, a SNP and an indel around 240bp apart, identified in the same 14 individuals of family 98002, had the greatest reduction on the linkage peak (Figure 5-6). Although their CADD scores do not predict deleterious functional consequences, the sites are evolutionarily conserved (Table 5-4), suggesting a selection pressure against changes at these sites. *CALN1* showed only minimal expression in the ciliary body and trabecular meshwork in our RNAseq database and the OTD plier scores were the lowest of the genes with OTD expression results (Table 5-5). The expression of *CALN1* is highest in the brain and it may play a role in the physiology of neurons [278], however, there is no link to eye physiology. Taken together, it is unlikely that *CALN1* itself is the functional gene driving the chromosome 7 linkage peak. Further sequencing in the 14 individuals of family 98002, who carry this variant which drives the linkage peak, may yield more promising results in finding a candidate gene for IOP regulation. The closest gene to *CALN1* associated with POAG is *ELN*, around 2Mb away, which was identified along with other novel loci in a large multi-ethnic GWAS [111], although this variant from the GWAS was not associated with IOP. It is possible that Family 98002 carry a variant in a novel IOP gene which is yet to be identified and further research is necessary to confirm this.

Two variants in the same nine members of family GTAS54 were prioritised in the chromosome 3 linkage interval. One of these variants is in *FAM86DP*, a pseudogene which codes for non-coding RNA. This gene is not highly expressed in the ciliary body or trabecular meshwork and its function is unknown. The other variant is in *PDZRN3*, a gene which is more highly expressed in the relevant ocular tissues than *FAM86DP* and many of the other genes assessed in this study (Table 5-4). *PDZRN3* may be involved in the development of blood vessels [279] and the differentiation of other cell types [280-282], but currently, a role for this gene in IOP or glaucoma has not been established. It is possible that these variants are on the same haplotype as the true functional variant, which has not been sequenced in the WES of this study. Further investigation with family GTAS54 is necessary to identify possible IOP candidate genes in this linkage region.

The chromosome 6 linkage region did not provide any prioritised variants or genes. This was the least significant linkage peak of all the identified linkage regions in this study. This region of chromosome 6 had poor sequence quality in this study and would need to be re-sequenced to confirm the linkage region identified here and to identify possible variants and genes of

interest. Currently it is premature to propose possible IOP candidate genes from this linkage peak.

Although this study has not conclusively identified variants with functional relevance involved in IOP regulation, linkage analysis has narrowed down the genome-wide search space to QTLs across six chromosomes. Potential candidate genes have been identified on chromosomes 2 and 15. Two interesting candidate genes, worthy of further consideration, have been proposed from the chromosome 2 QTL. *FMNL2* has already been associated with IOP and may be important in IOP regulation in Family 98002. *PKP4* is a novel IOP gene proposed in our study, with the identified SNP, rs148782148, itself a possible functional variant in family GTAS54. *HERC2*, on chromosome 15, is also proposed as a gene which may be involved in IOP regulation and worthy of further investigation.

Chapter 6
Discussion

The research presented in this thesis represents a novel contribution to our understanding of the genetic influences to POAG risk in large families. This study utilised whole exome sequencing data from large, extended families with clinical intermediate traits of glaucoma to identify variants involved with susceptibility to POAG. The families were large, multigenerational, enriched for POAG and were a valuable resource enabling this study to be conducted. In total, the five families consisted of over 500 family members with the older generations linking multiple branches of the later generations. Over 300 participants provided clinical data with 249 of these providing DNA samples, from which WES was conducted. This study analysed IOP and VCDR which are genetically correlated endophenotypes for POAG, as well as CCT, a POAG risk factor. The core analyses of this study were conducted blind to the POAG disease status which was only revealed after all analyses were completed.

The main focus was to identify novel variants and genes which influence POAG risk in the families. Before this could be undertaken, it was important to determine whether previously identified variants and loci affect the POAG risk in these families. A comprehensive database was established (Appendix B) with loci associated with POAG and its clinical intermediate traits identified from GWAS, candidate gene association studies and linkage studies. These loci were tested for association in the families using measured genotype association (MGA), a family-based association test conducted within a variance components framework using SOLAR. Significant association results were only obtained from within the linkage region loci, with no variants reaching significance from the previously associated SNPs or 20kb windows around these SNPs. Within the linkage regions, rs114429531 in *MST1* and three variants on a presumed haplotype in *FLRT3*, with large effect size and contribution to IOP variance, are worthy of further consideration. Overall, the published loci were not significantly associated with the clinical traits in these families. This analysis was only able to be conducted within WES regions due to the restricted coverage of whole exome sequencing.

With the recent advancement of polygenic risk score (PRS) analysis, it would be possible to genotype these family members for identified POAG or IOP risk loci and assess their genetic risk more broadly. Although PRS analysis is generally applied in case-control cohorts [81,

283], its use within families could be beneficial for identifying the common genetic risk background upon which the effects of rare variants might contribute [284]. In glaucoma, PRS analysis is a recent development [83-85] and has only been tested in case-control cohorts. The study by Craig et al. (2020) [83], tested for differences in the penetrance of the *MYOC* mutation, p.Gln368Ter (rs74315329), in different PRS groups and found that this mutation was more highly penetrant in those with high PRS scores, which has important implications for families carrying this mutation. Understanding the background PRS scores of the families in this study may enable a better understanding of the effects of the rare variants identified.

The research presented in this thesis focussed on identifying rare variants with large effect sizes affecting POAG endophenotypes, with the potential to be involved in POAG pathogenesis. Linkage analysis using the WES data was conducted to narrow down the genome search space and identify relevant variants in these families. Conducting linkage analysis using WES data is a more novel, but potentially more cost-effective strategy than conducting linkage using microsatellite or SNP array data followed by deeper sequencing [151, 219]. There are very few studies which have been published using linkage directly from exome sequencing data and these studies have been conducted on much smaller pedigrees than those of our study (for example [221, 234]). The use of such large pedigrees to conduct linkage analysis directly from exome sequencing data on the clinical intermediate traits of POAG was a novel approach.

One of the most important factors when conducting linkage analysis is to consider how to estimate identity by descent (IBD). The inheritance of alleles shared with a common ancestor is the key to linkage analysis. Our study chose to use the IBDLD program [152, 226], which was specifically designed for dense genotyping data, and was able estimate IBD between all the individuals in the pedigrees. Multipoint IBD matrices were generated and used for linkage analysis. Linkage disequilibrium (LD) is problematic in linkage analysis, as if it is not accounted for, there may be inflated LOD scores and false positive results [218, 285]. The IBDLD program incorporates LD into the estimation tool, rather than requiring LD pruning. Other studies conducting linkage directly from sequencing data have inferred genotypes at the locations of HapMap SNPs specifically to include common variants in linkage equilibrium [221, 234]. Those two studies involve much smaller pedigrees to those used in our study and it would be interesting to compare these two approaches to IBD estimation in large pedigrees.

Multipoint variance components linkage analysis was conducted using SOLAR and corrected for ascertainment of the POAG enriched families and the non-normal distribution of the IOP

and IOPmed traits, using an empirical LOD adjustment. Not only do these families have more POAG individuals than in the general population, but there are more individuals with extremely high IOP measures, causing the skewed distribution of IOP phenotype data compared to the general population. The VCDR measures were normally distributed and did not require ascertainment correction for the linkage analysis. The LOD adjustment conducted for IOPmed used a correction constant of 0.59, recognising the inflation of this trait in these families and the potential identification of false positive linkage regions. This provided confidence that the QTLs which were identified represented true linkage signals.

The analysis of the IOPmed trait identified six QTLs reaching at least suggestive significance (LOD score $\geq$ 1.86), on chromosomes 2, 3, 6, 7, 9 and 15. Two peaks, those on chromosomes 9 and 15, reached full significance (LOD score $\geq$ 3.3). All of the QTLs, other than that on chromosome 15, were novel for POAG and its endophenotypes. The chromosome 15 QTL identified here completely overlapped the GLC1I locus, with our linkage region and the published region displaying very similar shaped peaks [177]. The analysis of VCDR only resulted in one suggestive significant peak on chromosome 7. The lack of a significant QTL for this trait meant VCDR was not prioritised for further investigation in this study. Reanalysing this trait as a dichotomous trait, rather than quantitative, and trying different analysis approaches (as discussed in Chapter 5) which may be more suitable to this distribution of data, may improve the linkage signals for this trait within these families.

The IOPmed QTLs were analysed individually, by prioritising variants based on their presence in families contributing to the peak, their MGA results and then conducting linkage conditional on the measured genotype of the variants. Variants on the presumed haplotype as other prioritised variants were also included in downstream analysis. The QTLs on chromosomes 15 and 2 provided the most promising candidate variants and genes. A variant in *HERC2*, unobserved in the full gnomAD database (located at chr15: 28474713), is proposed as a novel variant influencing IOP in this study. This variant was identified in five members of Family 95002 and was associated with over 10mmHg increase in IOP and contributed to nearly 5% of the IOPmed variance. This is a very large increase in IOP when considered that the normal IOP is $\leq$ 21mmHg.

This is a private variant to Family 95002 and demonstrates the value in using families to identify rare variants with large effect sizes. Very recently, the *HERC2* locus was associated with POAG in a GWAS [83], and as the first reference of *HERC2* in the POAG literature, lends

additional weight to our findings. Two IOP genes are proposed from the chromosome 2 QTL, *PKP4* and *FMNL2*. *PKP4* is novel for IOP and POAG, with the identified variant (rs148782148) of potential functional relevance itself. This non-synonymous SNP was identified in 8 members of Family GTAS54 and was associated with an almost 5mmHg increase in IOP and nearly 6% of the IOPmed variance. Although not as rare as the *HERC2* variant, it has only been observed in 0.5% of the European population of the gnomAD database. *FMNL2* is proposed as the most likely functional gene for IOP on a haplotype identified in Family 98002. The same 19 members of this family carry variants in *RIF1*, *NEB* and *FMNL2*. *FMNL2* has already been associated with IOP and POAG [111, 171], and recently with VCDR [83]. Functional studies have proposed this gene's involvement with IOP, with a knockdown of *FMNL2* in a human trabecular meshwork cell line inducing morphological changes in the cells [111]. That research proposed that the effect on the contractile properties of trabecular meshwork cells could affect aqueous humour outflow, with a resultant increase in IOP as the mechanism of POAG pathogenesis.

As *FMNL2* variants had previously been identified in the literature, the question might be asked as to why the variant identified in this study (rs138926754) did not appear in Aim1 (Chapter 3), when published SNP loci and 20kb windows around SNP loci were examined for association in the families of this study. The variant in our study was in the 20kb window of one of the published variants (rs55692468) [171], however, it did not meet the MGA p-value nominal significance threshold (p-value = 0.19). Only nominally significant variants were discussed in Chapter 3. However, in Chapter 5, this *FMNL2* variant was identified on a presumed haplotype in the same individuals who also carried the *RIF1* variants, which were prioritised based on MGA p-value. The *FMNL2* SNP identified here is not proposed as the functional variant itself, but suggests the gene as a possible IOP regulating gene. Further sequencing and functional studies would need to be undertaken to identify possible functional variants, including those in regions not sequenced with WES, such as enhancers, long non-coding RNA and other non-coding regions.

The use of MGA served different purposes as described in Chapters 3 and 5. For Aim 1 (Chapter 3), this family-based association test was used to indicate whether previously published POAG loci impacted the clinical intermediate traits of POAG in the families of this study. For Aim 3 (Chapter 5) this test was used to prioritise variants in the newly identified linkage regions. In Aim 1, published POAG linkage regions spanning over 300 million base

pairs were tested for association using MGA and only 15 SNPs within 8 loci reached full statistical significance after correction for multiple testing. The most significant association identified was $p = 5.74x10^{-6}$ (for rs3755809 with CCT).

By narrowing the search space to the QTLs identified by linkage analysis, MGA could be used as a tool for prioritising variants within these regions, as demonstrated in Chapter 5. A nominal significance threshold of $p = 0.05$ was set for this analysis. Using the MGA results as a prioritisation tool, rather than prioritising using CADD scores and other *in silico* functional tools, retains the kinship information inherent in these families. This is an analysis designed for quantitative traits in families and it was important to retain as much familial information as possible. In diseases with a clear cut affected/unaffected diagnosis, the segregation of functional variants with disease is a first line prioritisation tool, which can then be followed by using CADD scores and other measures of deleteriousness. However, our study has used quantitative traits, rather than a discrete POAG diagnosis, and it was important to use the kinship information of the families to follow the inheritance of haplotypes which influence the proportion of trait variation in these families. These haplotypes can then be analysed further with targeted sequencing to identify variants with potential functional relevance. This is exemplified with the identification of the risk haplotype, spanning *RIF1*, *NEB* and *FMNL2* in the same 19 members of Family 98002. Some of the variants in these genes had MGA p-values above the nominal threshold, but reduced the linkage peak to similar levels as variants with p-values less than 0.05 when included in the linkage as covariates. So although MGA was a valuable tool in prioritising variants, it was necessary to also look beyond the p-values when selecting variants and genes of interest.

The identification of variants of interest relied on the reduction of the linkage peak when linkage was conducted conditional on the measured genotype of the variants being tested. If the linkage peak was reduced when the variant was included as a covariate, that variant was prioritised within that linkage region. One of the limitations with this approach is that this is not a sensitive method for testing each variant. Firstly, any variant in the same combination of individuals and with the same dosage of this allele, will produce the same results when included in the linkage model. This was observed, for example, in nine members of Family GTAS54 with the variants identified in *TTC21B* (rs572065013) and *DLX1* (rs186854626), which are located over 6 million bp apart on chromosome 2. This is likely to be the same haplotype in these individuals, although this would need to be validated.

Another limitation of linkage conditional on measured genotype, is that an additive model is assumed, with the heterozygous variant carrier halfway between the two homozygotes [246]. This model may not necessarily be appropriate for all variants; however, it is a practicable model when the true underlying model is unknown. Linkage conditional on measured genotype is a coarse method for identifying variants of interest which then need to be followed up with supporting evidence. Other factors need to be assessed, for example the measured genotype association results, including whether the variant accounts for much of the trait's variance and the effect on IOP (beta-value) as well as whether they are predicted to have functional consequences.

There are several strengths and limitations to be recognised in this body of research. A major strength is the families which were recruited into this study. They are large, both in depth and width, and enriched for glaucoma, with a higher rate of POAG than the general population. Recruitment of families such as these, with a complex, late-onset disease, is difficult. By the time the disease has manifested, many family members, such as parents and siblings, may already have died, while offspring may still be too young to manifest disease. To have access to both clinical and DNA data from so many individuals in these multigenerational families has provided a strong foundation for this research. Although clinical and DNA data were predominantly available for the younger members, these families were linked by comprehensive and well curated genealogical information, which provided more power to estimate trait heritabilities and to conduct linkage analysis [286]. Over 300 of these family members provided clinical data, however, CCT measures were only obtained from 186 of these (148 also with sequencing data), due to these measurements commencing routinely later in the study, limiting the power to use this trait for linkage analysis. The highly significant heritabilities for all of the traits obtained in these families, provided the power to conduct the association and linkage analyses. However, the p-values for trait heritabilities were not significant when families were modelled individually. Hence, to determine family contribution to the QTLs and retain power (significant heritability p-values), four families were tested at a time, with one family excluded in turn.

The use of quantitative traits was a strength of this study. A POAG diagnosis can be subjective, when a certain threshold of vision loss and nerve damage is reached. But the use of objective quantitative traits, which can be measured in individuals with and without POAG, provides more power for linkage and association [42, 220, 287]. The fact that these clinical

measurements were taken over multiple time-points for most of the family members in this study improves the chance that more representative measures for these traits were used, rather than just a single measurement at one time-point (described in Chapter 2.3). The use of intermediate traits to study complex disease is now well recognised [37, 42, 68, 127] and provides a strategy to simplify the genetic architecture. The use of a variance components framework for the linkage analysis allowed a decomposition of the total trait variance into QTL specific variances and provided confidence that the linkage signals for IOPmed were independent and not the result of false signals.

The WES sequencing data provided both strengths and limitations to this study. The strength in using WES sequencing rather than microsatellite markers or SNP arrays is that variants of potential functional significance within coding regions can be identified from the linkage regions, without the need for further sequencing. The *PKP4* variant within the chromosome 2 linkage region is a potentially functional variant involved in IOP regulation, which was identified directly from the QTL. However, most of the variants which were prioritised within the linkage peaks were not necessarily good functional candidates themselves. The identified variants may be tagging the true functional variants on the same haplotype. Either these variants were not sequenced in the WES, or they were sequenced, but did not meet quality thresholds (80% high confidence calls). Although the sequencing quality may have affected a few variants, we believe that many of the true functional variants lie in non-coding regions of the genome and were not sequenced in this study, as evidenced by the predominance of significant non-coding hits in recent GWAS (for example, [83, 85, 111]).

The importance of non-coding variants in complex disease is becoming better understood as research has increased in this area.. In POAG, the most replicated locus is *CDKN2B-AS1*, a long non-coding RNA [91, 99] (Appendix B). In a study on another highly replicated POAG locus, a variant in the enhancer of *SIX6* resulted in increased expression of the gene and dysregulation of protein expression [108]. In a study assessing variants in miRNA coding genes in POAG, an association was found with a common variant in *MIR182*, particularly in the high-tension glaucoma subset [288]. In another study using data extracted from a GWAS in which variants were identified in miRNAs as well as 3'UTRs in miRNA binding sites, associations were identified with two miRNA variants and VCDR and cup area, as well as associations between 47 3'UTR sites with VCDR, IOP and cup area [289]. That study exemplifies not only that SNPs identified in GWAS may represent tagging of the true functional variant in LD but

that the GWAS SNPs themselves may be the true functional variants in non-coding regions [290]. If the variants identified in our study are not the true functional variants themselves, targeted sequencing of the candidate gene is required to identify more plausible candidates.

Following on from the research presented in this thesis, there are several areas for further investigation. The first task is to use Sanger sequencing [291] to validate the variants identified in *HERC2*, *PKP4* and *FMNL2* (from Aim 3) as well as the *MST1* and *FLRT3* variants (from Aim 1). Once accurate genotyping has been confirmed, screening these variants, and other variants in the same genes, in large case/control cohorts, for example the Australian and New Zealand Registry of Advanced Glaucoma (ANZRAG) cohort [292] may identify differences in allele frequency between cases and controls and indicate promising variants and genes. This is exemplified in a study which screened GWAS identified POAG and endophenotype genes in the WES sequencing data of the ANZRAG cohort and compared it with sequencing data from controls [147]. Rare, predicted pathogenic variants within *CARD10*, a gene associated with VCDR and optic disc parameters, were found to be enriched in the POAG cohort. Testing for enrichment of the variants and other variants in the genes identified in our study in the ANZRAG cohort may be able to prioritise genes for further analysis. Targeted sequencing of genes identified in the families of this study may also lead to the identification of more promising variants, including non-coding variants. Promising candidate genes could be taken into cell lines of appropriate ocular tissues, such as trabecular meshwork, ciliary body, and retinal ganglion cells and knocked down or mutated to investigate their effects. The *FMNL2* functional research discussed above, involving gene silencing, was conducted in a human trabecular meshwork cell line [111]. Although retinal ganglion cells are difficult to grow in culture, promising research has been conducted on *OPTN* mutations inducing cell death in a cell line expressing markers of retinal ganglion cells [293, 294].

The WES data generated for our study provided rich information which can be used further to enhance the research presented here. Copy number variants (CNVs) were not assessed in our study, and with recent advances in tools for detecting CNVs [295], may prove to be a valuable pathway to follow. The *TBK1* CNV is well recognised in causing 1% of NTG cases [57, 61, 62]. Other, as yet unidentified CNVs may be involved with the clinical intermediate traits of POAG in the families of this study. The WES data could also be used as a scaffold to impute data in non-coding regions of the genome. Linkage analysis could then be conducted on the expanded dataset, which may lead to the identification of novel peaks for IOP and/or VCDR.

Any identified peaks would need targeted sequencing to identify variants of interest. As WES data is only rarely used in linkage analysis, imputing from WES data prior to linkage analysis would provide an innovative strategy to search for linkage regions in non-coding regions. Although not a linkage analysis methodology, imputation from WES data was successful in identifying uncommon variants involved with quantitative traits of blood cell disorders [296].

Another strategy for enhancing the WES data of this study would be to phase the genotypes and search for compound heterozygous variants in these families. *CYP1B1* is a gene important in the pathogenesis of glaucoma [65-67] and compound heterozygotes have been associated with more severe forms of POAG [297], JOAG [298] and PCG [299]. Identifying compound heterozygous variants would be novel for the clinical intermediate traits of POAG and worthy of further investigation.

The research conducted in this study was based on a linkage analysis methodology to identify rare variants. Non-linkage based methods are also suitable for rare variant identification with the WES data presented here. Software such as; RVTESTS [300], famSKAT [301], gskat [302] and MONSTER [303] were designed to detect association of rare variants with quantitative traits in family based data. Recently, RVTESTS was used, in conjunction with other software, to identify novel loci associated with IOP, VCDR and POAG [304]. Although that study was a meta-analysis of GWAS data, and not family based, RVTESTS was also designed to detect association of rare variants in related individuals [300]. The use of such software, with the WES data from our study, will provide a complementary approach to the linkage analysis presented in this dissertation, in searching for rare variants associated with POAG.

The ultimate goal of genetics research on POAG is to improve the diagnosis and treatment for this complex disease. Considering that 50% of Australians currently living with POAG are undiagnosed [6], improved screening for this disease is necessary to prevent permanent vision loss. Genetic screening has the potential to identify at risk individuals, preferably before vision loss occurs. With our current understanding of glaucoma genetics, with a few rare causative mutations along with hundreds of variants each associated with a slightly increased risk, there is no justification for genetic testing of POAG at the population level [305]. However, for at risk populations, such as family members of those with POAG, or those with high IOP, polygenic risk score analysis is providing potential for predicting the chances of developing this disease, the severity and even the need for surgery [83-85, 248]. For those families carrying a *MYOC* mutation, there is much promise with predictive genetic testing, identifying at risk

family members at an earlier stage than those who presented through clinical pathways [54, 80]. Potentially, if the variants or genes identified in our study prove to be of functional importance in regulating IOP, genetic testing could be offered to the members of that family, allowing for more regular monitoring of IOP and possibly an earlier intervention with pressure reducing medication.

Finding novel genes involved with IOP as well as confirming genes already identified, can only improve our understanding of the biological pathways involved with the pathogenesis of POAG. This study has proposed a novel gene for IOP, *PKP4*, as well as a gene already identified as involved with IOP, *FMNL2*. In the gene interaction analysis, there were identified protein interactions between CDH1 with both FMNL2 and PKP4. All three of these proteins are involved with cell-cell adhesion complexes [111, 253, 274], however, PKP4 is novel to IOP and glaucoma. The interaction of PKP4 with CDH1, which itself has an interaction with FMNL2 and which are both recognised IOP genes, is a promising finding. The novel *HERC2* variant is also a promising finding, not only because of its associated effect on IOP in the variant carriers, but this locus has also recently been associated with POAG [83]. Although no mechanisms of involvement were suggested in that study, the research conducted here may propose an IOP mediated pathway.

A better understanding of the genetics of POAG and the biological pathways in which these genes act, will allow for a greater variety of treatments in the future [78]. Exciting developments using CRISPER-Cas9 technology [306] have the potential for targeted and permanent genetic therapies. For example, mutant *MYOC* genes have been knocked down in human trabecular meshwork cells and in a mouse model to reduce IOP, and has the potential to revolutionise treatment for those carrying *MYOC* mutations [307]. A less personalised approach, but perhaps more beneficial to a larger group of individuals, is the promising research on using the same technology to knock down the aquaporin 1 gene (*AQP1*), involved with aqueous humour secretion from the ciliary body [308]. Although variants in this gene were tested but found not to be associated with POAG [309], this gene is believed to be involved with actin fibre reorganisation with possible effects to trabecular meshwork cell contractility [310]. *FMNL2* is also believed to affect actin fibres in trabecular meshwork cells [111] and future research may possibly find a biological link between these two genes. The aquaporin knockdown study [308] demonstrated successful trials in mouse models and human tissue culture and propose that a single injection into the eye can target the ciliary body to

permanently reduce IOP. As more genes involved with POAG and it endophenotypes are uncovered, a better understanding of the pathogenesis of this complex disease will be developed, along with more progressive and personalised therapies.

Conclusion

This novel study has investigated the genetic susceptibility to POAG by searching for rare variants with large effect sizes in the whole exome sequencing data of extended families. Several rare variants associated with IOP increase were identified, demonstrating enrichment in the families of this study and validating the use of extended pedigrees in complex disease research. These rare variants may not yet be useful targets for population-wide screening as POAG risk alleles, but they are useful within the families themselves, allowing family members who are carriers earlier and more frequent IOP monitoring. The variants identified in this study propose novel genes for involvement with IOP; *MST1*, *FLRT3*, *PKP4* and *HERC2*, whilst also proposing *FMNL2* which had previously been associated with IOP and POAG. The recent association of the *HERC2* locus with POAG in an independent study, in parallel with our study, gives confidence that the rare variant identified here is a promising finding.

Appendices

Appendix A General Methods Scripts

A.1 Configuration file for bcbio variant calling

Template for whole genome Illumina variant calling with GATK pipeline

Saved as config_file-template.yaml

```
      details:
        - analysis: variant2
          genome_build: hg19
```

to do multi-sample variant calling, assign samples the same metadata / batch
metadata:
batch: allvariants

```
          algorithm:
            aligner: false
            recalibrate: false
            bam_clean: remove_extracontigs
            variantcaller: gatk-haplotype
            jointcaller: gatk-haplotype-joint
```

for targeted projects, set the region

```
variant_regions: ~/TruSeq-Exome-Targeted-Regions.bed
      resources:
        samtools:
          cores: 1
        sambamba:
          cores: 1
```

A.2 Vcf annotations

A.2.1 Script to add GATK filtering tags to genotype calls

```
java -Xmx4g -jar /gd/apps/GATK/GenomeAnalysisTK.jar \
-R /gd/apps/gatk-bundle/hg19/ucsc.hg19.fasta \
-T VariantFiltration \
-o output_file_name.vcf \
--variant input_file_name.vcf \
--genotypeFilterExpression "DP >= 10 && GQ >= 20" --genotypeFilterName
"High_Confidence" \
--genotypeFilterExpression "DP < 10 && GQ < 20" --genotypeFilterName
"Low_Confidence" \
--genotypeFilterExpression "DP < 10 && GQ >= 20" --genotypeFilterName
"Low_coverage_High_quality" \
--genotypeFilterExpression "DP >= 10 && GQ < 20" --genotypeFilterName
"High_coverage_Low_quality"
```

A.2.2 ANNOVAR annotation of vcf file

```
table_annovar.pl input_file_name.vcf \
/gd/apps/annovar-2018Apr16/humandb -buildver hg19 \
-out input_file_name_annotated.vcf \
-remove -otherinfo \
-protocol refGene,avsnp150,popfreq_max_20150413,\
1000g2015aug_all,1000g2015aug_eur,1000g2015aug_afr,1000g2015aug_amr,\
1000g2015aug_eas,1000g2015aug_sas,esp6500siv2_all,esp6500siv2_ea,\
esp6500siv2_aa,exac03,exac03nontcga,exac03nonpsych,kaviar_20150923,\
gnomad_exome,gnomad_genome,clinvar_20170905,cadd13,gerp++elem,\
dbnsfp31a_interpro,spidex,dbscsnv11,dbnsfp33a,gwava,eigen,fathmm \
-operation g,f,f,f,f,f,f,f,f,f,f,f,f,f,f,f,f,f,f,f,f,f,f,f,f,f,f \
-nastring . -vcfinput
```

A.3 R script converting bcbio called vcf to dosage file for MGA

Script to convert bcbio SNPs and indels vcf files (un-annotated) to dosage file suitable for running MGA in SOLAR

All columns need names.

```
library(tidyr)
library(dplyr)
```

Remove the # before the CHROM in the heading, or it won't be read in

```
POAG <- read.table(file = "input_file_name.vcf",
                   sep="\t",header=TRUE,as.is=TRUE,na.strings=".",
                   stringsAsFactors = FALSE, colClasses = "character")
```

Convert ./. reads to blank cells. Note the . needs to be "escaped" and the . at the end will match any character.

```
POAG <- as.data.frame(sapply(POAG,gsub,pattern="[.]/[.].*",replacement=""))
```

To filter variants to get rid of the really rare ones.

Count the number of heterozygous variant calls each variant in the file and add this as a new column in the vcf file

The ^ is to only count the genotypes at the beginning of the each cell (the info column has "unwanted" genotypes in it)

The [/|] is to pick up both unphased and phased genotypes. Prob won't make a difference here, but might for future runs

```
POAG$het_counts <- apply(POAG, 1, function(x)
                   sum(grepl('^[1-9][/|]0|^0[/|][1-9]', x)))
```

Count the number of homozygote variant calls each variant in the file and add this as a new column in the vcf file

```
POAG$hom_counts <- apply(POAG, 1, function(x)
                   sum(grepl('^[1-9][/|][1-9]', x)))
```

Create another column which adds 1 per het count and 2 per homozygous count

```
POAG$total_counts <- (POAG$het_counts + 2*POAG$hom_counts)
```

Filter vcf file to get rid of rare or too common SNPs (filter is a dplyr function, so make sure it's loaded)

```
filtered_between <- filter(POAG,
                     (POAG$total_counts >= 5 & POAG$total_counts <= 494))
```

After filtering, delete the 3 columns which I added in to do the counts (columns 259 to 261)

```
filtered_POAG <- filtered_between[, -c(259:261)]
```

Replace remaining variant strings with counts

the ^ ensures that the genotype at the start of the cell is replaced.

In bcbio files, there are annotations in the info column which would get recognised too.

```
filtered_POAG <-
as.data.frame(sapply(filtered_POAG,gsub,pattern="^0/0.*",replacement="0"))
filtered_POAG <-
as.data.frame(sapply(filtered_POAG,gsub,pattern="^1/1.*",replacement="2"))
```

But some indels (not many though) are multiallelic. So anything which is 0/* or */0 will be a 1 dosage:

```
filtered_POAG <- as.data.frame(sapply(filtered_POAG,gsub,pattern="^0/[1-
9].*",replacement="1"))
filtered_POAG <- as.data.frame(sapply(filtered_POAG,gsub,pattern="^[1-
9]/0.*",replacement="1"))
```

And for the homozygous multiallelics

```
filtered_POAG <- as.data.frame(sapply(filtered_POAG,gsub,pattern="^[1-
9]/[1-9].*",replacement="2"))
```

Make the dosage file of all variants

adds a new ID column into the data.frame which is what will be used for the variant ID for SOLAR

```
filtered_POAG$New_ID <- paste(filtered_POAG$CHROM, filtered_POAG$POS,
sep="_")
```

Subsets out just the genotypes (dosages) for the samples,

```
POAG_subset <- subset(filtered_POAG, select = 10:258)
```

Make the row names of the subset be the new variant idenifier made above, prefixed by snp_ which is required in SOLAR

```
rownames(POAG_subset) <- paste("snp_",filtered_POAG$New_ID, sep="")
```

Flip the data.frame around so samples are now rows instead of columns and genotypes are columns instead of rows

```
transpose_POAG_subset <- as.data.frame(t(POAG_subset))
```

remove the X from the row names

```
row.names(transpose_POAG_subset) <-
gsub("X","",row.names(transpose_POAG_subset))
```

Quality control step

we work with a matrix (because nchar not vectorized for data.frame)

nchar returns 0 for "" and 1 for "0" "1" or "2"

```
test <- nchar(as.matrix(transpose_POAG_subset))
```

plot the sums so we get a sense

```
plot(colSums(test))
```

create a logical vector to identify columns to keep

colSums is a vector of counts of TRUE based on test above

```
keep <- colSums(test) >= (nrow(test) * 1/2)
```

finally apply our identifier to extract the columns to keep (≥ 50% dosages per variant)

```
clean <- transpose_POAG_subset[, keep]
```

write out to csv for SOLAR input

```
write.table(clean, file = "trait_dosages_date.csv", quote = FALSE, sep =
",", na = "NA", row.names = TRUE, col.names = TRUE)
```

This csv file needs the SNP columns moved one to the right and ID labelling the first column

A.4 Measured genotype association analysis in SOLAR

load pedigree and phenotype files

full pedigree and phenotype files used (531 family members)

```
load ped pedigree_file.ped
load phen phenotype_file.phen


model new
trait TRAIT
```

covariates to be included

abbreviation of interactions of age and age2 with sex

```
covar age^1,2#sex study
```

to run variance components modelling, only including the significant covariates

```
polygenic -screen
```

to run measured genotype association analysis

```
mgassoc -file dosage_file
```

A.5 Example linkage diagram R script

This script is to create plots with multiple lines superimposed and a legend on the bottom.

For all the chromosomes in a grid pattern

To load ggplot and tidyr

```
library("ggplot2")
library("tidyr")
```

Read in file (5 columns - Chromosome, Location 3 different LODs)

```
trait <- read.table("table_with_multipoint_results.txt", \
header = TRUE, sep = "\t")
```

Adjust data into "long format"

```
trait_long <- gather (trait, IOP, LOD, 3:5, -Chromosome)
```

Draw graphs - 1 box for each autosome. Comparing 3 traits in the same graph

```
ggplot(data=trait_long,
      aes(x=Location, y=LOD, colour=IOP, linetype=IOP)) +
      theme(legend.position = c(0.45, 0.05),
      legend.title = element_blank(),
      axis.title.x = element_text(hjust=0.18)) + xlab('Location (cM)') +
      facet_wrap( ~ Chromosome, switch = "x") +
      scale_colour_manual(values = c('blue','magenta', 'green4')) +
      scale_linetype_manual(values = c(1,1,1)) +
      geom_line(size = 0.5) +
      geom_hline(yintercept=3, lty="solid", colour = "grey50, size=0.5") +
      #geom_hline(yintercept=2, lty="dashed", colour = "grey50", size =0.5)
```

Appendix B POAG loci database

Key - column headings:

SNP/indel = single nucleotide polymorphism / insertion or deletion

MIM number = genotype or phenotype number from the Online Mendelian Inheritance in Man catalogue (http://omim.org/)

Locus = location of the nearest gene

Chr = chromosome number

Chr. position (bp) = base pair position of each SNP in the two most recent human genome assemblies as found in the dbSNP database (http://www.ncbi.nlm.nih.gov/SNP/).

POAG = primary open-angle glaucoma

IOP = intraocular pressure

VCDR = vertical cup to disc ratio

CCT = central corneal thickness

ODA = optic disc area

OCA = optic cup area

Key – body of table

1 = association found between SNP and POAG or endophenotype.

0 = no association found between SNP and POAG or endophenotype.

J = juvenile open angle glaucoma

N = normal tension glaucoma

C = control cohort

D = discovery cohort

R = replication cohort

NTG = normal tension glaucoma cohort

HTG = high tension glaucoma cohort

AS = association study using a candidate gene approach

GWAS = genome-wide association study

GWAS (Ped.) = genome-wide association study using pedigrees

LS = linkage study

MA = meta-analysis

Population codes:

AA	African-American
AC	Afro-Caribbean population of Barbados, West Indies
AF	Africa
AS	Asia
AU	Australia
AT	Austria
BMES	Blue Mountains Eye Study
BR	Brazil
CA	Canada
Cau	Caucasian
CH	China
CR	Croatia
EG	Egypt
ES	Estonia
EU	Europe
FR	France
GE	Germany
HI	Hispanic
IC	Iceland
IGGC	International Glaucoma Genetics Consortium
IN	India
IR	Iran
IT	Italy
JA	Japan
KO	Korea
LA	Latino
ME	Mediterranean
NIES	Norfolk Island Eye Study
NL	Netherlands
NZ	New Zealand
PA	Pakistan
PH	Philipines
PO	Poland
SA	Saudi Arabia
SC	Singaporean-Chinese
SCL	Scotland
SM	Singaporean-Malay
SW	Sweden
TU	Turkey
UK	United Kingdom
US	United States

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	Locus/Gene/ Nearest gene (SNPs)	Gene name	MIM number	Locus	Chr	Hg19 (2009)	Hg38 (2013)	POAG	IOP	VCDR	CCT	ODA or OCA	Population	Study Type	Study size	Reference
indel	*PRDM16*	PR domain containing 16	605557	1p36.32	1	3,046,430	3,129,866			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs12028027	*PRDM16*	PR domain containing 16	605557	1p36.32	1	3,050,331	3,133,767			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs2252865	*RERE*	RE repeats-encoding gene	605226	1p36.23	1	8,422,676	8,362,616	1		1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs301801	*RERE*	RE repeats-encoding gene	605226	1p36.23	1	8,495,945	8,435,885			1			EU, AS	GWAS	27,878	Springelkamp et al., 2014 [103]
rs1801133	*MTHFR*	5,10-methylenetetra-hydrofolate reductase	607093	1p36.3	1	11,856,378	11,796,321	1					GE	AS	218	Junemann et al., 2005 [311]
rs1801133	*MTHFR*	5,10-methylenetetra-hydrofolate reductase	607093	1p36.3	1	11,856,378	11,796,321	0					IR	AS	248	Nilforoushan et al., 2012 [312]
rs1801133	*MTHFR*	5,10-methylenetetra-hydrofolate reductase	607093	1p36.3	1	11,856,378	11,796,321	1					SA	AS	210	Al-Shahrani et al., 2016
promoter	*NPPA*	natriuretic peptide precursor A	108780	1p36.22	1	11,907,741 - 11,908,341	11,847,684 - 11,848,284	1					AU	AS	113	Tunny et al., 1996 [181]
rs873458	*MFN2*	mitofusin 2	608507	1p36.2	1	12,046,089	11,986,032	N					GE	AS	567	Wolf et al., 2009 [313]
rs2295281	*MFN2*	mitofusin 2	608507	1p36.2	1	12,059,412	11,999,355	N					GE	AS	567	Wolf et al., 2009 [313]
rs11588779	*MFN2*	mitofusin 2	608507	1p36.2	1	12,082,881	12,022,824	N					GE	AS	567	Wolf et al., 2009 [313]
rs3924048	*DHRS3*	short-chain dehydrogenase / reductase family, member 3	612830	1p36.21	1	12,614,848	12,554,818			1			EU AS	GWAS	17248 6841	Springelkamp et al., 2015 [134]
rs274754	*COL8A2*	collagen, type VIII, alpha-2	120252	1p34.2	1	36,565,617	36,100,016				1		US-Cau	AS	100	Desronvil et al., 2010 [314]
rs96067	*COL8A2*	collagen, type VIII, alpha-2	120252	1p34.2	1	36,571,920	36,106,319				1		US-Cau	AS	100	Desronvil et al., 2010 [314]
rs96067	*COL8A2*	collagen, type VIII, alpha-2	120252	1p34.2	1	36,571,920	36,106,319				1		D1: SM D2: SG-CH	GWAS	D1: 3280 D2: 3400	Vithana et al., 2011 [315]
rs96067	*COL8A2*	collagen, type VIII, alpha-2	120252	1p34.2	1	36,571,920	36,106,319				1		D: AS R: AS	GWAS	D: 7711 R: 2681	Cornes et al., 2012 [316]
rs96067	*COL8A2*	collagen, type VIII, alpha-2	120252	1p34.2	1	36,571,920	36,106,319				0		D: GE R: NL	GWAS	D: 3931 R: 1418	Hoehn et al., 2012 [198]
rs96067	*COL8A2*	collagen, type VIII, alpha-2	120252	1p34.2	1	36,571,920	36,106,319				0		D: US-Cau AS: US-Cau	GWAS AS	D: 1117 AS: 6469	Ulmer et al., 2012 [196]
rs96067	*COL8A2*	collagen, type VIII, alpha-2	120252	1p34.2	1	36,571,920	36,106,319				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
rs96067	*COL8A2*	collagen, type VIII, alpha-2	120252	1p34.2	1	36,571,920	36,106,319				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs1866758	HIVEP3	human immunodeficiency virus type 1 enhancer-binding protein 3	606649	1p34.2	1	42,318,868	41,853,197		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	**Locus/Gene/ Nearest gene (SNPs)**	**Gene name**	**MIM number**	**Locus**	**Chr**	**Hg19 (2009)**	**Hg38 (2013)**	**POAG**	**IOP**	**VCDR**	**CCT**	**ODA or OCA**	**Population**	**Study Type**	**Study size**	**Reference**
	–			1p32	1	50,700,001 - 61,300,000	50,200,001 - 60,800,000			2			AU	LS	1 large family	Charlesworth et al., 2005 [189]
rs1925953	*RPE65*	retinoid isomerohydrolase	180069	1p31.3	1	68,848,681	68,382,998	1		1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs1192414	*CDC7/TGFBR3*	cell division cycle 7/transforming growth factor, beta receptor III	603311/ 600742	1p22.1	1	92,075,134	91,609,577	1		1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs1192415	*CDC7/TGFBR3*	cell division cycle 7/transforming growth factor, beta receptor III	603311/ 600742	1p22.1	1	92,077,097	91,611,540					1	D:NL R:NL, UK	GWAS	D: 7360 R: 4455	Ramdas et al., 2010 [99]
rs1192415	*CDC7/TGFBR3*	cell division cycle 7/transforming growth factor, beta receptor III	603311/ 600742	1p22.1	1	92,077,097	91,611,540	1					EU	AS	45998	Ramdas et al., 2011 [69]
rs1192415	*CDC7/TGFBR3*	cell division cycle 7/transforming growth factor, beta receptor III	603311/ 600742	1p22.1	1	92,077,097	91,611,540					1	D:AS R:NL	GWAS	D: 4445 R: 9326	Khor et al., 2011 [317]
rs1192415	*CDC7/TGFBR3*	cell division cycle 7/transforming growth factor, beta receptor III	603311/ 600742	1p22.1	1	92,077,097	91,611,540					1	D:NL	GWAS	D: 23000	Axenovitch et al., 2011 [104]
rs1192415	*CDC7/TGFBR3*	cell division cycle 7/transforming growth factor, beta receptor III	603311/ 600742	1p22.1	1	92,077,097	91,611,540	0					AF	AS	437	Cao et al., 2012 [96]
rs1192415	*CDC7/TGFBR3*	cell division cycle 7/transforming growth factor, beta receptor III	603311/ 600742	1p22.1	1	92,077,097	91,611,540					1	AU, NZ	AS	1759	Dimasi et al., 2012 [29]
rs1192415	*CDC7/TGFBR3*	cell division cycle 7/transforming growth factor, beta receptor III	603311/ 600742	1p22.1	1	92,077,097	91,611,540	1					AS, AF, EU	GWAS	D: 13250 R: 35953	Li et al., 2015 [100]
rs1192415	*CDC7/TGFBR3*	cell division cycle 7/transforming growth factor, beta receptor III	603311/ 600742	1p22.1	1	92,077,097	91,611,540			1	1		IN	AS	468	Philomenadin et al., 2015 [105]
rs1192415	*CDC7/TGFBR3*	cell division cycle 7/transforming growth factor, beta receptor III	603311/ 600742	1p22.1	1	92,077,097	91,611,540					1	EU AS	GWAS	17248 6841	Springelkamp et al., 2015 [134]
rs1192415	*CDC7/TGFBR3*	cell division cycle 7/transforming growth factor, beta receptor III	603311/ 600742	1p22.1	1	92,077,097	91,611,540	1					US, L, AS, AA	AS	63,412	Choquet et al., 2018[111]
rs4658101	*CDC7/TGFBR3*	cell division cycle 7/transforming growth factor, beta receptor III	603311/ 600742	1p22.1	1	92,077,409	91,611,852			1			EU, AS	GWAS	27,878	Springelkamp et al., 2014 [103]
rs4658101	*CDC7/TGFBR*	cell division cycle 7/transforming growth factor, beta receptor III	603311/ 600742	1p22.2/1 p22.1	1	92,077,409	91,611,852	1		1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs148639588	*COL11A1*	collagen, type XI, alpha 1	120280	1p21.1	1	103,272,424	102,806,868	1				1	EU	GWAS	OAG:3071 C:6750	Gharahkhani et al., 2018 [131]
deletion	*GSTM1*	glutathione S-transferase, mu-1	138350	1p13.3	1	110,230,418 - 110,236,367	109,687,795 - 109,693,744	1					ES	AS	452	Juronen et al., 2000 [318]
deletion	*GSTM1*	glutathione S-transferase, mu-1	138350	1p13.3	1	110,230,418 - 110,236,367	109,687,795 - 109,693,744	1					IT	AS	87	Izzotti et al., 2003 [319]
deletion	*GSTM1*	glutathione S-transferase, mu-1	138350	1p13.3	1	110,230,418 - 110,236,367	109,687,795 - 109,693,744	0					SW	AS	400	Jansson et al., 2003 [320]
deletion	*GSTM1*	glutathione S-transferase, mu-1	138350	1p13.3	1	110,230,418 - 110,236,367	109,687,795 - 109,693,744	1					TU	AS	312	Yildirim et al., 2005 [321]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	**Locus/Gene/ Nearest gene (SNPs)**	**Gene name**	**MIM number**	**Locus**	**Chr**	**Hg19 (2009)**	**Hg38 (2013)**	**POAG**	**IOP**	**VCDR**	**CCT**	**ODA or OCA**	**Population**	**Study Type**	**Study size**	**Reference**
deletion	*GSTM1*	glutathione S-transferase, mu-1	138350	1p13.3	1	110,230,418 - 110,236,367	109,687,795 - 109,693,744	1					SA	AS	169	Abu-Amero et al., 2008 [323]
deletion	*GSTM1*	glutathione S-transferase, mu-1	138350	1p13.3	1	110,230,418 - 110,236,367	109,687,795 - 109,693,744	0					CH	AS	606	Fan et al., 2010 [264]
deletion	*GSTM1*	glutathione S-transferase, mu-1	138350	1p13.3	1	110,230,418 - 110,236,367	109,687,795 - 109,693,744	1					BR	AS	172	Rocha et al., 2011 [324]
rs856077	*IFI16*	interferon-gamma-inducible protein 16	147586	1q23.1	1	158,969,075	158,999,285				1		NIES	GWAS (Ped.)	330	Matovinovic et al., 2017 [133]
rs4656461	*TMCO1*	transmembrane and coiled-coil domains 1	614123	1q24	1	165,687,205	165,717,968	1					D: AU, NZ R: AU, NZ	GWAS	D: 4546 R: 4148	Burdon et al., 2011 [91]
rs4656461	*TMCO1*	transmembrane and coiled-coil domains 1	614123	1q24	1	165,687,205	165,717,968	1	1	0	0		AU, NZ	AS	1420	Sharma et al., 2012 [109]
rs4656461	*TMCO1*	transmembrane and coiled-coil domains 1	614123	1q24	1	165,687,205	165,717,968	1					D: EU R: EU	GWAS	D: 3077 R: 13044	Gharahkhani et al., 2014 [89]
rs4656461	*TMCO1*	transmembrane and coiled-coil domains 1	614123	1q24	1	165,687,205	165,717,968	1	1				D: AS D: EU	GWAS	D: 7738 D:27558	Hysi et al., 2014 [86]
rs4656461	*TMCO1*	transmembrane and coiled-coil domains 1	614123	1q24	1	165,687,205	165,717,968	1					BMES	AS	OAG: 67 C: 1919	Burdon et al 2015 [95]
rs4656461	*TMCO1*	transmembrane and coiled-coil domains 1	614123	1q24	1	165,687,205	165,717,968	1					US, L, AS, AA	AS	63,412	Choquet et al., 2018 [111]
rs7555523	*TMCO1*	transmembrane and coiled-coil domains 1	614123	1q24	1	165,687,205	165,717,968	1	1				D: AS D: EU	GWAS	D: 7738 D:27558	Hysi et al., 2014 [86]
rs6668108	TMCO1	transmembrane and coiled-coil domains 1	614123	1q24.1	1	165,691,320	165,722,083	1	1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs7524755	TMCO1	transmembrane and coiled-coil domains 1	614123	1q24.1	1	165,694,897	165,725,660	1	1				D: US, L, AS, AA R: multi ethnic	GWAS	D: 63,412 R: 176,890	Choquet et al., 2018 [111]
rs10918274	*TMCO1*	transmembrane and coiled-coil domains 1	614123	1q24	1	165,714,416	165,745,179	1	1				EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs7555523	*TMCO1*	transmembrane and coiled-coil domains 1	614123	1q24	1	165,718,979	165,749,742	1	1				D: NL R: UK, AU, NZ,CA AS: NL, GE	GWAS AS	D: 11972 R: 7482 AS: 1432	van Koolwijk et al., 2012 [66]
rs7518099	*TMCO1*	transmembrane and coiled-coil domains 1	614123	1q24	1	165,736,880	165,767,643	1					D: AU, NZ R: AU, NZ	GWAS	D: 4546 R: 4148	Burdon et al., 2011 [91]
rs7518099	*TMCO1*	transmembrane and coiled-coil domains 1	614123	1q24	1	165,736,880	165,767,643	0					D: UK R:UK	GWAS	D: 437 R: 344	Gibson et al., 2012 [93]
rs7518099	*TMCO1*	transmembrane and coiled-coil domains 1	614123	1q24	1	165,736,880	165,767,643		1				D: AS D: EU	GWAS	D: 7738 D:27558	Hysi et al., 2014 [86]
rs7518099	*TMCO1*	transmembrane and coiled-coil domains 1	614123	1q24	1	165,736,880	165,767,643		1				D: EU	GWAS	D: 6236	Ozel et al., 2014 [110]
rs7518099	*TMCO1*	transmembrane and coiled-coil domains 1	614123	1q24	1	165,736,880	165,767,643	1					Meta-analysis US, AU, EU, SC	GWAS	37333	Bailey et al., 2016 [12]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	**Locus/Gene/ Nearest gene (SNPs)**	**Gene name**	**MIM number**	**Locus**	**Chr**	**Hg19 (2009)**	**Hg38 (2013)**	**POAG**	**IOP**	**VCDR**	**CCT**	**ODA or OCA**	**Population**	**Study Type**	**Study size**	**Reference**
rs2814471	*TMCO1*	transmembrane and coiled-coil domains 1	614123	1q24	1	165,739,598	165,770,361	1					D: multi ethnic R: US, L, AS, AA	GWAS	D: 176,890 R: 63,412	Choquet et al., 2018 [111]
rs2239854	*F5*	coagulation factor V	612309	1q24.2	1	169,525,808	169,556,570			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
indel	*F5*	coagulation factor V	612309	1q24.2	1	169,530,520	169,561,282			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs12406092	*F5*	coagulation factor V	612309	1q24.2	1	169,543,131	169,573,893					1	EU AS	GWAS	17248 6841	Springelkamp et al., 2015 [134]
rs10753787	*F5*	coagulation factor V	612309	1q24.2	1	169,549,775	169,580,537			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
	*MYOC/TIGR (*GLC1A*)*	myocilin/trabecular meshwork-induced glucocorticoid response protein	601652	1q24.3	1	171,604,557 - 171,621,773	171,635,416 - 171,652,632	1 J					US	LS	1 family	Locus: Sheffield et al., 1993 [46] Gene: Stone et al., 1997 [55]
	OPTC	opticin	605127	1q32.1	1			1					IN	AS	300	Acharya et al., 2007 [325]
rs4846476	*TGFB2*	transforming growth factor beta 2	190220	1q41	1	218,526,228	218,352,886				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs596169	MIR548F3	microRNA 548f-3		1q41	1	219,147,419	218,974,077		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
c.109 G>A; p.Val37Met	*TP53BP2*	tumor protein p53 binding protein 2	602143	1q41	1	223,967,595 - 224,033,674	223,779,893 - 223,845,972	1					NL	LS	1 family	Michael, et al., 2017 [146]
rs6671926	*CDC42BPA*	CDC42 binding protein kinase alpha	603412	1q42.13	1	227,386,971	227,199,270					1	EU AS	GWAS	17248 6841	Springelkamp et al., 2015 [134]
rs547984	*ZP4*	zona pellucida glycoprotein 4	613514	1q43	1	238,096,886	237,933,586	1					D: JA R: JA	GWAS	D: 1519 R: 700	Nakano et al., 2009 [76]
rs547984	*ZP4*	zona pellucida glycoprotein 4	613514	1q43	1	238,096,886	237,933,586	0					AF	AS	437	Cao et al., 2012 [96]
rs540782	*ZP4*	zona pellucida glycoprotein 4	613514	1q43	1	238,097,039	237,933,739	1					D: JA R: JA	GWAS	D: 1519 R: 700	Nakano et al., 2009 [76]
rs693421	*ZP4*	zona pellucida glycoprotein 4	613514	1q43	1	238,099,090	237,935,790	1					D: JA R: JA	GWAS	D: 1519 R: 700	Nakano et al., 2009 [76]
rs693421	*ZP4*	zona pellucida glycoprotein 4	613514	1q43	1	238,099,090	237,935,790	0					CH	AS	1039	Chen et al., 2012b [326]
rs693421	*ZP4*	zona pellucida glycoprotein 4	613514	1q43	1	238,099,090	237,935,790	1					KO	AS	1115	Kim et al., 2014 [327]
rs2499601	*ZP4*	zona pellucida glycoprotein 4	613514	1q43	1	238,104,895	237,941,595	1					D: JA R: JA	GWAS	D: 1519 R: 700	Nakano et al., 2009 [76]
rs13016883	*TRIB2*	tribbles pseudokinase 2	609462	2p24.3	2	12,877,307	12,737,181	1		1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs2113818	*TRIB2*	tribbles pseudokinase 2	609462	2p24.3	2	12,890,860	12,750,734			1			EU AS	GWAS	17248 6841	Springelkamp et al., 2015 [134]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	Locus/Gene/ Nearest gene (SNPs)	Gene name	MIM number	Locus	Chr	Hg19 (2009)	Hg38 (2013)	POAG	IOP	VCDR	CCT	ODA or OCA	Population	Study Type	Study size	Reference
rs115781177	*LTBP1*	latent transforming growth factor beta binding protein 1	150390	2p22.3	2	33,348,494	33,123,427				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs115179432	LTBP1	latent transforming growth factor-beta-binding protein 1	150390	2p22.3	2	33,348,679	33,123,612		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
	GLC3A *(CYP1B1)*		601771 (231300)	2p22.2	2	38,294,746 - 38,303,323	38,067,602 - 38,076,180	1					TU	LS	17 families	Sarfarazi et al., 1995 [50]
rs1800440	*CYP1B1*	cytochrome P450, subfamily 1, polypeptide 1	601771	2p22.2	2	38,298,139	38,070,996	1					PA	AS	330	Micheal et al., 2015[182]
rs1800440	*CYP1B1*	cytochrome P450, subfamily 1, polypeptide 1	601771	2p22.2	2	38,298,139	38,070,996	1					Mixed	MA	8 studies	Wang et al., 2015 [183]
rs1056827	*CYP1B1*	cytochrome P450, subfamily 1, polypeptide 1	601771	2p22.2	2	38,302,177	38,075,034	1					PA	AS	330	Micheal et al., 2015 [182]
rs1056827	*CYP1B1*	cytochrome P450, subfamily 1, polypeptide 1	601771	2p22.2	2	38,302,177	38,075,034	1					Mixed	MA	8 studies	Wang et al., 2015 [183]
rs10012	*CYP1B1*	cytochrome P450, subfamily 1, polypeptide 1	601771	2p22.2	2	38,302,390	38,075,247	1					PA	AS	330	Micheal et al., 2015 [182]
rs2617266	*CYP1B1*	cytochrome P450, subfamily 1, polypeptide 1	601771	2p22.2	2	38,302,544	38,075,401	1					PA	AS	330	Micheal et al., 2015 [182]
rs3213787	*SRBD1*	S1 RNA binding domain 1	—	2p21	2	45,646,824	45,419,685	N					D: JA	GWAS	D: 660	Meguro et al., 2010 [328]
rs3213787	*SRBD1*	S1 RNA binding domain 1	—	2p21	2	45,646,824	45,419,685	1					JA	AS	561	Mabuchi et al., 2011 [329]
rs3213787	*SRBD1*	S1 RNA binding domain 1	—	2p21	2	45,646,824	45,419,685	0					AF	AS	437	Cao et al., 2012 [96]
rs11884064	*SRBD1*	S1 RNA binding domain 1	—	2p21	2	45,837,053	45,609,914	1					D: UK R: UK	GWAS	D: 437 R: 344	Gibson et al., 2012 [93]
	GLC1H		611276	2p16–2p15	2	47,800,001 - 64,100,000	47,500,000 - 63,900,000	1					AC, UK	LS	several families	Suriyapperuma et al., 2007 [176]
rs1533428	*LOC730100*	uncharacterized LOC730100	—	2p16.3	2	51,959,258	51,732,120	1					CH	AS	1039	Chen et al., 2012b [326]
rs7426380	*EFEMP1*	EGF containing fibulin-like extracellular matrix protein 1	601548	2p16.1	2	56,095,010	55,867,875		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs3791679	*EFEMP1*	EGF-containing fibuliin-like extracellular matrix protein 1	601548	2p16.1	2	56,096,892	55,869,757			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs1346786	*EFEMP1*	EGF-containing fibuliin-like extracellular matrix protein 1	601548	2p16.1	2	56,108,333	55,881,198			1			EU AS	GWAS	17248 6841	Springelkamp et al., 2015 [134]
	–			2p14	2	68,139,016-68,339,157	67,911,884-68,112,025	1					US	LS	sibpairs	Wiggs et al., 2000 [190]
rs6732795	*ANTXR1*	anthrax toxin receptor 1	606410	2p13.3	2	69,411,517	69,184,385		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs678350	*HK2*	hexokinase 2	601125	2p12	2	75,062,870	74,835,743	1 N					JA	AS	1108	Shi et al., 2013 [330]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	**Locus/Gene/ Nearest gene (SNPs)**	**Gene name**	**MIM number**	**Locus**	**Chr**	**Hg19 (2009)**	**Hg38 (2013)**	**POAG**	**IOP**	**VCDR**	**CCT**	**ODA or OCA**	**Population**	**Study Type**	**Study size**	**Reference**
	GLC1B		606689	2cen–2q13	2	92,326,172 - 114,400,000	93,900,000 - 112,200,000	1						LS	Several families	Stoilova et al., 1996 [175]
rs2033008	*NCK2*	NCK adaptor protein 2	604930	2q12	2	106,502,585	105,886,129	N					JA	AS	246	Akiyama et al., 2008 [331]
rs2033008	*NCK2*	NCK adaptor protein 2	604930	2q12	2	106,502,585	105,886,129	N					JA	AS	1108	Shi et al., 2013 [330]
rs1800587	*IL1A*	interluekin 1 alpha	147760	2q14	2	113,542,960	112,785,383	1					CH	AS	323	Wang et al., 2006a [184]
rs1143634	*IL1B*	interluekin 1 beta	147720	2q14	2	113,590,390	112,832,813	1					CH	AS	163	Lin et al., 2003b [332]
rs7588567	*NCKAP5*	NCK-associated protein 5	608789	2q21.2	2	134,363,032	133,605,461	1					D:JA R:JA	GWAS	D: 7993 R: 9014	Osman et al., 2012 [90]
rs7588567	*NCKAP5*	NCK-associated protein 5	608789	2q21.2	2	134,363,032	133,605,461	0					US, L, AS, AA	AS	63,412	Choquet et al., 2018 [111]
rs56117902	*FMNL2*	formin-like 2	616285	2q23.3	2	153,304,730	152,448,216	1	1				D: US, L, AS, AA R: multi ethnic	GWAS	D: 63,412 R: 176,890	Choquet et al., 2018 [111]
rs55692468	*FMNL2*	formin-like 2	616285	2q23.3	2	153,361,375	152,504,861		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
	–			2q33.1–2q33.3	2	197,400,001 - 209,000,000	196,600,001 - 208,200,000	1					AC	LS	several families	Nemesure et al., 2003 [187]
rs56335522	*IKZF2*	IKAROS family zinc finger 2	606234	2q34	2	213,758,234	212,893,510	1					D: multi ethnic R: US, L, AS, AA	GWAS	D: 176,890 R: 63,412	Choquet et al., 2018 [111]
rs1549733	*DIRC3*	disrupted in renal carcinoma 3	608262	2q35	2	218,472,172	217,607,449					1	EU AS	GWAS	17248 6841	Springelkamp et al., 2015 [134]
rs1035673	*TNS1*	tensin 1	600076	2q35	2	218,675,533	217,810,810		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs10189064	*USP37*	ubiquitin specific peptidase 37		2q35	2	219,327,500	218,462,777				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
rs121908120	*WNT10A*	wingless-type MMTV integration site family, member 10A	606268	2q35	2	219,755,011	218,890,289				1		D: EU R: EU	EWAS	D:1029 R:4479	Cuellar-Partida et al., 2015 [333]
rs121908120	*WNT10A*	wingless-type MMTV integration site family, member 10A	606268	2q35	2	219,755,011	218,890,289				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs4608502	*COL4A3*	collagen, type IV, alpha-3	120070	2q36.3	2	228,134,155	227,269,439				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs7606754	*COL4A3*	collagen, type IV, alpha-3	120070	2q36.3	2	228,135,180	227,270,464				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
rs143937055	*COL4A3*	collagen, type IV, alpha-3	120070	2q36.3	2	228,143,967	227,279,251		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs7599762	*COL6A3*	collagen, type VI, alpha-3	120250	2q37.3	2	238322885	237,414,242		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs1801282	*PPARG*	peroxisome proliferator activated receptor gamma	601487	3p25.2	3	12,393,125	12,351,626	1					IN	AS	122	Chandra et al., 2016 [334]
rs690037	*RTFN1*	raftlin, lipid raft linker 1		3p24	3	16,395,668	16,354,161					1	D:AU R:UK	GWAS	D: 1368 R: 848	Macgregor et al. 2010 [335]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	Locus/Gene/ Nearest gene (SNPs)	Gene name	MIM number	Locus	Chr	Hg19 (2009)	Hg38 (2013)	POAG	IOP	VCDR	CCT	ODA or OCA	Population	Study Type	Study size	Reference
rs11129176	*RARB*	retinoic acid receptor, beta	180220	3p24.2	3	25,049,310	25,007,819					1	EU AS	GWAS	17248 6841	Springelkamp et al., 2015 [134]
rs11129176	*RARB*	retinoic acid receptor, beta	180220	3p24.2	3	25,049,310	25,007,819			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
	GLC1L		137750	3p22–3p21	3	38,048,763 - 54,479,766	38,007,272 - 54,445,739	1					AU	LS	1 family	Baird et al., 2005 [174]
rs2710323	*ITIH1*	inter-alpha-trypsin inhibitor, heavy chain 1	147270	3p21.1	3	52,815,905	52,781,889	1					D: EU R: EU	GWAS	D: 3077 R: 13044	Gharahkhani et al., 2014 [89]
rs6764184	*FLNB*	filamin b	603381	3p14.3	3	58,006,266	58,020,539			1			EU AS	GWAS	17248 6841	Springelkamp et al., 2015 [134]
rs6764184	*FLNB*	filamin b	603381	3p14.3	3	58,006,266	58,020,539	1		1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs12494328	*FLNB*	filamin b	603381	3p14.3	3	58,035,497	58,049,770	1		1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs13093086	*FOXP1*	forkhead box P1	605515	3p13	3	71,055,162	71,006,011		1				D: EU	GWAS	D: 6236	Ozel et al., 2014 [110]
rs34201102	*CADM2*	cell adhesion molecule 2	609938	3p12.1	3	85,137,499	85,088,348	1					D: multi ethnic R: US, L, AS, AA	GWAS	D: 176,890 R: 63,412	Choquet et al., 2018 [111]
rs3749260	*GPR15*	G protein-coupled receptor 15	601166	3q11.2	3	98,250,862	98,532,018				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
rs2623325	*COL8A1*	collagen, type VIII, alpha-1	120251	3q12.1	3	99,131,755	99,412,911			1			EU AS	GWAS	27,878	Springelkamp et al., 2014 [103]
rs6804624	*COL8A1*	collagen, type VIII, alpha-1	120251	3q12.1	3	99,159,147	99,440,303			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs1997404	*COL8A1*	collagen, type VIII, alpha-1	120251	3q12.1	3	99,161,022	99,442,178			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs9860250	*ABI3BP*	ABI family, member 3 (NESH) binding protein	606279	3q12.2	3	100,637,871	100,919,027					1	EU AS	GWAS	17248 6841	Springelkamp et al., 2015 [134]
rs9843102	*ABI3BP*	ABI family, member 3 (NESH) binding protein	606279	3q12.2	3	100,650,929	100,932,085	1		1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
	GLC1C		601682	3q21 – 3q24	3	121,900,001 - 148,900,000	122,200,001 - 149,200,000	1					US	LS	1 family	Wirtz et al., 1997 [153]
indel	*STAG1*	stromal antigen 1	604358	3q22.3	3	136,138,073	136,419,231				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs367923973	*IL20RB*	interleukin 20 receptor beta	605621	3q22.3	3	136,701,097	136,982,255	1					US	AS/ Gene	Family	Wirtz & Keller, 2016 [336]
rs1164313	*SLC9A9*	solute carrier family 9 member A9	608396	3q24	3	144,339,407	144,620,565				1		NIES	GWAS (Ped.)	330	Matovinovic et al., 2017 [133]
rs13095933	*AGTR1*	angiotensin receptor type 1	106165	3q24	3	148,408,108	148,690,321				1		NIES	GWAS (Ped.)	330	Matovinovic et al., 2017 [133]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	Locus/Gene/ Nearest gene (SNPs)	Gene name	MIM number	Locus	Chr	Hg19 (2009)	Hg38 (2013)	POAG	IOP	VCDR	CCT	ODA or OCA	Population	Study Type	Study size	Reference
rs11710845	*LINC01214 / TSC22D2*	long intergenic non-protein coding RNA 1214 / TSC22 domain family, member 2	?/617724	3q25.1	3	150,065,280	150,347,493		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs344002	*TIPARP / LINC00886*	TCDD-inducible poly(ADP-ribose) polymerase / Long intergenic non-protein coding RNA 886	612480/-	3q25.31	3	156,455,269	156,737,480				1		NIES	GWAS (Ped.)	330	Matovinovic et al., 2017 [133]
rs9822953	*TIPARP*	TCDD-inducible poly (ADP-ribose) polymerase	612480	3q25.31	3	156,472,071	156,754,282				1		EU, AS	GWAS	>20,000	Lu et al., 2013 [200]
rs9822953	*TIPARP*	TCDD-inducible poly (ADP-ribose) polymerase	612480	3q25.31	3	156,472,071	156,754,282				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs7636836	*FNDC3B*	fibronectin domain III-containing protein 3B	611909	3q26.31	3	171,765,125	172,047,335	1					Japan	GWAS	POAG:7378 C:36385	Shiga et al., 2018 [337]
rs4894796	*FNDC3B*	fibronectin domain III-containing protein 3B	611909	3q26.31	3	171,780,763	172,062,973	1					AS, AF, EU	GWAS	D: 13250 R: 35953	Li et al., 2015 [100]
rs6445046	*FNDC3B*	fibronectin domain III-containing protein 3B	611909	3q26.31	3	171,933,252	172,215,462				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs7635832	*FNDC3B*	fibronectin domain III-containing protein 3B	611909	3q26.31	3	171,989,276	172,271,486		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs7635832	*FNDC3B*	fibronectin domain III-containing protein 3B	611909	3q26.31	3	171,989,276	172,271,486	1	1				EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs6445055	*FNDC3B*	fibronectin domain III-containing protein 3B	611909	3q26.31	3	171,992,387	172,274,597		1				D: AS D: EU	GWAS	D: 7738 D:27558	Hysi et al., 2014 [86]
rs4894535	*FNDC3B*	fibronectin domain III-containing protein 3B	611909	3q26.31	3	171,995,605	172,277,815				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
rs7620503	*TBL1XR1*	transducin-beta-like 1 receptor 1	608628	3q26.32	3	177,304,298	177,586,510				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
indel	*TBL1XR1*	transducin-beta-like 1 receptor 1	608628	3q26.32	3	177,306,757	177,588,969				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs2111534	*MFN1*	mitofusin 1	608506	3q25–3q26	3	179,090,772	179,372,984	N					GE	AS	567	Wolf et al., 2009 [313]
rs1000002	*PARL/ABCC5*	presenilin associated, rhomboid like protein / ATP-binding cassette, subfamily C, member 5	607858 /605251	3q27	3	183,635,768	183,917,980	N					GE	AS	567	Wolf et al., 2009 [313]
rs1402003	*PARL/ABCC5*	presenilin associated, rhomboid like protein / ATP-binding cassette, subfamily C, member 5	607858 /605251	3q27	3	183,642,879	183,925,091	N					GE	AS	567	Wolf et al., 2009 [313]
rs9853115	*DGKG*	diacylglycerol kinase, gamma	601854	3q27.3	3	186,131,600	186,413,811		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs9853115	*DGKG*	diacylglycerol kinase, gamma	601854	3q27.3	3	186,131,600	186,413,811	1					D: multi ethnic R: US, L, AS, AA	GWAS	D: 176,890 R: 63,412	Choquet et al., 2018 [111]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	Locus/Gene/ Nearest gene (SNPs)	Gene name	MIM number	Locus	Chr	Hg19 (2009)	Hg38 (2013)	POAG	IOP	VCDR	CCT	ODA or OCA	Population	Study Type	Study size	Reference
rs56962872	*LINC02052 / CRYGS*	long intergenic non-protein coding RNA 2052 / crystallin, gamma-S	-/123730	3q27.3	3	186,200,223	186,482,434	1		1			EU	GWAS	OAG:3071 C:6750	Gharahkhani et al., 2018 [131]
rs13076750	*LPP*	LIM domain-containing preferred translocation partner in lipoma	600700	3q28	3	188,059,443	188,341,655		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs166850	*OPA1*	optic atrophy 1	605290	3q29	3	193,355,074	193,637,285	N					UK	AS	276	Aung et al., 2002 [338]
rs166850	*OPA1*	optic atrophy 1	605290	3q29	3	193,355,074	193,637,285	0					Korea	AS	166	Woo et al., 2004 [339]
rs166850	*OPA1*	optic atrophy 1	605290	3q29	3	193,355,074	193,637,285	1					AF	AS	157	Yao et al., 2006 [340]
rs166850	*OPA1*	optic atrophy 1	605290	3q29	3	193,355,074	193,637,285	0					GE	AS	567	Wolf et al., 2009 [313]
rs166850	*OPA1*	optic atrophy 1	605290	3q29	3	193,355,074	193,637,285	0					CH	AS	606	Fan et al., 2010 [264]
rs166850	*OPA1*	optic atrophy 1	605290	3q29	3	193,355,074	193,637,285	1					UK	AS	212	Yu-Wai-Man et al., 2010 [341]
rs10451941	*OPA1*	optic atrophy 1	605290	3q29	3	193,355,102	193,637,313	N					UK	AS	276	Aung et al., 2002 [338]
rs10451941	*OPA1*	optic atrophy 1	605290	3q29	3	193,355,102	193,637,313	1					UK	AS	110	Powell et al., 2003 [342]
rs10451941	*OPA1*	optic atrophy 1	605290	3q29	3	193,355,102	193,637,313	0					KO	AS	166	Woo et al., 2004 [339]
rs10451941	*OPA1*	optic atrophy 1	605290	3q29	3	193,355,102	193,637,313	1					AF	AS	157	Yao et al., 2006 [340]
rs10451941	*OPA1*	optic atrophy 1	605290	3q29	3	193,355,102	193,637,313	1					JA	AS	570	Mabuchi et al., 2007 [343]
rs10451941	*OPA1*	optic atrophy 1	605290	3q29	3	193,355,102	193,637,313	0					GE	AS	567	Wolf et al., 2009 [313]
rs10451941	*OPA1*	optic atrophy 1	605290	3q29	3	193,355,102	193,637,313	0					CH	AS	606	Fan et al., 2010 [264]
rs10451941	*OPA1*	optic atrophy 1	605290	3q29	3	193,355,102	193,637,313	1					UK	AS	212	Yu-Wai-Man et al., 2010 [341]
rs6855176	*AFAP1*	actin filament-associated protein 1	608252	4p16.1	4	7,753,403	7,751,676	1					Japan	GWAS	POAG:7378 C:36385	Shiga et al., 2018 [337]
rs4619890	*AFAP1*	actin filament-associated protein 1	608252	4p16.1	4	7,853,160	7,851,433	1					D: EU R: EU	GWAS	D: 3077 R: 13044	Gharahkhani et al., 2014 [89]
rs4619890	*AFAP1*	actin filament-associated protein 1	608252	4p16.1	4	7,853,160	7,851,433	1					US, L, AS, AA	AS	63,412	Choquet et al., 2018 [111]
rs9330348	*AFAP1*	actin filament-associated protein 1	608252	4p16.1	4	7,883,887	7,882,160	1					D: multi ethnic R: US, L, AS, AA	GWAS	D: 176,890 R: 63,412	Choquet et al., 2018
rs28795989	*AFAP1*	actin filament-associated protein 1	608252	4p16.1	4	7,891,545	7,889,818	1	1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs59521811	*AFAP1*	actin filament-associated protein 1	608252	4p16.1	4	7,909,772	7,908,045	1	1				D: US, L, AS, AA R: multi ethnic	GWAS	D: 63,412 R: 176,890	Choquet et al., 2018 [111]
rs11732100	*AFAP1*	actin filament-associated protein 1	608252	4p16.1	4	7,924,690	7,922,963	1					Meta-analysis US, AU, EU, SC	GWAS	37333	Bailey et al., 2016 [12]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	**Locus/Gene/ Nearest gene (SNPs)**	**Gene name**	**MIM number**	**Locus**	**Chr**	**Hg19 (2009)**	**Hg38 (2013)**	**POAG**	**IOP**	**VCDR**	**CCT**	**ODA or OCA**	**Population**	**Study Type**	**Study size**	**Reference**
rs17527016	*PITX2*	Paired-like homeodomain transcription factor 2	601542	4q25	4	111,963,719	111,042,563		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs17017968	*FREM3*	FRAS1-related extracellular matrix protein 3	608946	4q31.21	4	144,542,213	143,621,060		1				US-Cau	GWAS	D: 1660	Chen et al., 2015 [143]
rs5335	*EDNRA*	Endothelin receptor, type A	131243	4q31.2	4	148,463,840	147,542,688	N					JA	AS	650	Ishikawa et al., 2005 [344]
rs28789690	*NR3C2*	Nuclear receptor subfamily 3, group C, member 2	600983	4q31.23	4	149,077,899	148,156,748				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs3931397	*NR3C2*	Nuclear receptor subfamily 3, group C, member 2	600983	4q31.23	4	149,079,497	148,158,346				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
	GLC1Q		–	4q35.1– 4q35.2	4	183,200,001 - 191,154,276	182,278,848 - 190,123,121	1					UK	LS	1 family	Porter et al., 2011 [345]
rs76325372	*ANKH*	ankylosis, homolog (mouse)	605145	5p15.2	5	14,837,332	14,837,223	1					D: multi ethnic R: US, L, AS, AA	GWAS	D: 176,890 R: 63,412	Choquet et al., 2018 [111]
rs72759609	*PDZD2*	PDZ domain-containing 2	610697	5p13.3	5	31,952,051	31,951,945	1		1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs4865762	*MOCS2/FST*	molybdenum cofactor synthesis gene 2 / Follistatin	603708 /136470	5q11.2	5	52,582,931	53,287,101		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs61275591	*ANKRD55*	ankyrin repeat domain 55	615189	5q11.2	5	55,775,556	56,479,729	1					Japan	GWAS	POAG:7378 C:36385	Shiga et al., 2018 [337]
rs1309531	*CWC27*	CWC27 spliceosome associated protein homolog	617170	5q12.3	5	64,306,311	65,010,484				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs1117707	*CWC27/ ADAMTS6*	CWC27 spliceosome associated protein homolog / A disintegrin-like and metalloproteinase with thrombospondin type 1 motif, 6	-/605008	5q12.3	5	64,389,665	65,093,838				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
rs10471310	*ADAMTS6*	a disintegrin-like and metalloproteinase with thrombospondin type 1 motif, 6	605008	5q12.3	5	64,548,961	65,253,134				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs2307121	*ADAMTS6*	a disintegrin-like and metalloproteinase with thrombospondin type 1 motif, 6	605008	5q12.3	5	64,625,512	65,329,685				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
rs10064391	*ADAMTS6*	a disintegrin-like and metalloproteinase with thrombospondin type 1 motif, 6	605008	5q12.3	5	64,686,659	65,390,832				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs7717697	*VCAN*	versican	118661	5q14.2- 5q14.3	5	82,744,604	83,448,785			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs73220188	*FBXL17/FER*	F-box and leucine-rich repeat protein 17 / FER tyrosine kinase	609083 /176942	5q21.3	5	108,053,612	108,717,911		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	**Locus/Gene/ Nearest gene (SNPs)**	**Gene name**	**MIM number**	**Locus**	**Chr**	**Hg19 (2009)**	**Hg38 (2013)**	**POAG**	**IOP**	**VCDR**	**CCT**	**ODA or OCA**	**Population**	**Study Type**	**Study size**	**Reference**
	GLC1M		610535	5q22.1-5q32	5	109,600,001 - 149,800,000	110,200,001 - 150,400,000	J					PH	LS	1 family	Pang et al., 2006 [48]
	WDR36 (GLC1G)	WD repeat-containg protein 36	609669	5q22.1	5	110,427,870 - 110,466,200	111,092,172 - 111,130,502	1 N					US	LS	1 large family	Locus and gene: Monemi et al., 2005 [73]
rs249767	*FGF1*	fibroblast growth factor 1	131220	5q31.3	5	141,918,585	142,539,020				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs1042714	*ADRB2*	beta-2-adrenergic receptor	109690	5q32–5q34	5	148,206,473	148,826,910		1				JA	Gene	745	Inagaki et al., 2006b [186]
rs17658229	*DUSP1*	dual-specificity phosphatase 1	600714	5q35.1	5	172,191,052	172,764,049			1			EU, AS	GWAS	27,878	Springelkamp et al., 2014 [103]
rs35084382	*DUSP1*	dual-specificity phosphatase 1	600714	5q35.1	5	172,197,039	172,770,036			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs35028368*	*ADAMTS2*	a disintegrin-like and metalloproteinase with thrombospondin type 1 motif, 2	604539	5q35.3	5	178,671,146	179,244,145				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs17756712	*EXOC2*	exocyst complex component 2	615329	6p25.3	6	625,071	625,071			1			EU, AS	GWAS	27,878	Springelkamp et al., 2014 [103]
rs2073006	*EXOC2*	exocyst complex component 2	615329	6p25.3	6	637,465	637,465	1					D: multi ethnic R: US, L, AS, AA	GWAS	D: 176,890 R: 63,412	Choquet et al., 2018
rs2745572	*FOXC1*	forkhead box C1	601090	6p25.3	6	1,548,369	1,548,134	1					Meta-analysis US, AU, EU, SC	GWAS	37333	Bailey et al., 2016 [12]
rs2745572	*FOXC1*	forkhead box C1	601090	6p25.3	6	1,548,369	1,548,134	1					US, L, AS, AA	AS	63,412	Choquet et al., 2018 [111]
rs2745572	*FOXF2/ FOXCUT*	forkhead box F2 / FOXC1 upstream transcript, noncoding	603250 / 615976	6p25.3	6	1,548,369	1,548,134	1	1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs11969985	*GMDS*	GDP-mannose 4,6-dehydratase	602884	6p25.3	6	1,922,907	1,922,673	1					D: EU R: EU	GWAS	D: 3077 R: 13044	Gharahkhani et al., 2014 [89]
rs11969985	*GMDS*	GDP-mannose 4,6-dehydratase	602884	6p25.3	6	1,922,907	1,922,673	1					US, L, AS, AA	AS	63,412	Choquet et al., 2018
rs4960295	*RREB1*	ras responsive element binding protein 1	602209	6p24.3	6	7,205,796	7,205,563			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs4645836	*TNFα*	tumour necrosis factor	191160	6p21.33	6	31,542,476	31,574,699	1					CH	AS	464	Wang et al., 2012 [346]
rs1800629	*TNFα*	tumour necrosis factor	191160	6p21.33	6	31,543,031	31,575,254	1					CH	AS	163	Lin et al., 2003 [347]
rs1800629	*TNFα*	tumour necrosis factor	191160	6p21.33	6	31,543,031	31,575,254	0					AU	AS	342	Mossbock et al., 2006 [348]
rs1800629	*TNFα*	tumour necrosis factor	191160	6p21.33	6	31,543,031	31,575,254	1					Iran	AS	289	Razeghinejad et al., 2009 [349]
rs1800629	*TNFα*	tumour necrosis factor	191160	6p21.33	6	31,543,031	31,575,254	1					CH	AS	606	Fan et al., 2010 [264]
rs1800629	*TNFα*	tumour necrosis factor	191160	6p21.33	6	31,543,031	31,575,254	1					TU	AS	279	Bozkurt et al., 2012 [350]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	**Locus/Gene/ Nearest gene (SNPs)**	**Gene name**	**MIM number**	**Locus**	**Chr**	**Hg19 (2009)**	**Hg38 (2013)**	**POAG**	**IOP**	**VCDR**	**CCT**	**ODA or OCA**	**Population**	**Study Type**	**Study size**	**Reference**
rs1800629	*TNFα*	tumour necrosis factor	191160	6p21.33	6	31,543,031	31,575,254	1					CH	AS	464	Wang et al., 2012 [346]
rs1043618	*HSPA1A*	Heat shock 70-KD protein 1A	140550	6p21.33	6	31,783,507	31,815,730			1			PO	AS	769	Nowak et al., 2015 [351]
exon 4	*TAP1*	transporter, ATP-binding cassette, major histocompatibility complex 1	170260	6p21.32	6	32,818,721 - 32,818,926	32,850,944 - 32,851,149	1					CH	AS	171	Lin et al., 2004 [352]
exon 10	*TAP1*	transporter, ATP-binding cassette, major histocompatibility complex 1	170260	6p21.32	6	32,814,844 - 32,814,981	32,847,068 - 32,847,204	1					CH	AS	171	Lin et al., 2004 [352]
rs72852338	*CDKN1A*	cyclin dependent kinase inhibitor 1A	116899	6p21.2	6	36,554,240	36,586,463			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs67530707*	*CDKN1A*	cyclin dependent kinase inhibitor 1A	116899	6p21.2	6	36,592,986	36,625,209	1		1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs67530707*	*CDKN1A*	cyclin dependent kinase inhibitor 1A	116899	6p21.2	6	36,592,986	36,625,209	1					US, L, AS, AA	AS	63,412	Choquet et al., 2018
rs6913530	*CDKN1A*	cyclin dependent kinase inhibitor 1A	116899	6p21.2	6	36,598,209	36,630,432	1					D: US, L, AS, AA R: multi ethnic	GWAS	D: 63,412 R: 176,890	Choquet et al., 2018
rs1801270	*CDKN1A*	cyclin dependent kinase inhibitor 1A	116899	6p21.2	6	36,651,971	36,684,194	1					CH	AS	118	Tsai et al., 2004 [353]
rs13191376	*RUNX2*	runt-related transcription factor 2	600211	6p21.1	6	45,522,139	45,554,402				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs1396046	*PKHD1*	polycystic kidney and hepatic disease 1 gene	606702	6p12.3 - p12.2	6	51,536,992	51,672,194		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs2025751	*PKHD1*	polycystic kidney and hepatic disease 1 gene	606702	6p12.3	6	51,622,449	51,757,651		1				D: EU	GWAS	D: 6236	Ozel et al., 2014 [110]
rs735860	*ELOVL5*	elongation of very long chain fatty acids-like 5	611805	6p12.1	6	53,123,118	53,258,320	N					D: JA	GWAS	D: 660	Meguro et al., 2010 [328]
rs1412710	*COL12A1*	collagen, type XII, alpha-1	120320	6q13-q14	6	75,837,203	75,127,487				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs1931656	*FAM46A*	family with sequence similarity 46, member a	611357	6q14.1	6	82,610,188	81,900,471				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs2875087	*TENT5A/IBTK*	terminal nucleotidyltransferase 5A / inhibitor of Bruton agammaglobulinemia tyrosine kinase	611357 /606457	6q14.1	6	82,616,083	81,906,366		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs9361886	*IBTK*	inhibitor of Bruton agammaglobulinemia tyrosine kinase	606457	6q14.1	6	82,778,502	82,068,785				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs1538138	*IBTK*	inhibitor of Bruton agammaglobulinemia tyrosine kinase	606457	6q14.1	6	82,794,594	82,084,877				1		D: AS R: AS	GWAS	D: 7711 R: 2681	Cornes et al., 2012 [316]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	**Locus/Gene/ Nearest gene (SNPs)**	**Gene name**	**MIM number**	**Locus**	**Chr**	**Hg19 (2009)**	**Hg38 (2013)**	**POAG**	**IOP**	**VCDR**	**CCT**	**ODA or OCA**	**Population**	**Study Type**	**Study size**	**Reference**
rs1538138	*IBTK*	inhibitor of Bruton agammaglobulinemia tyrosine kinase	606457	6q14.1	6	82,794,594	82,084,877				1		D: US-Cau AS: US-Cau	GWAS AS	D: 1117 AS: 6469	Ulmer et al., 2012 [196]
rs1538138	*IBTK*	inhibitor of Bruton agammaglobulinemia tyrosine kinase	606457	6q14.1	6	82,794,594	82,084,877				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
rs141917145*	*RFPL4B*	ret finger protein-like 4B		6q21	6	113,375,847	113,054,645		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs1402538	*HSF2*	heat shock transcription factor 2	140581	6q22.31	6	122,388,851	122,067,705			1			EU AS	GWAS	17248 6841	Springelkamp et al., 2015 [134]
rs868153	*HSF2*	heat shock transcription factor 2	140581	6q22.31	6	122,389,955	122,068,809			1			EU, AS	GWAS	27,878	Springelkamp et al., 2014 [103]
rs2684249	*HSF2*	heat shock transcription factor 2	140581	6q22.31	6	122,392,511	122,071,365	1		1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs9494457	*PDE7B*	phosphodiesterase 7B	604645	6q23.3	6	136,474,794	136,153,656	1					D: US, L, AS, AA R: multi ethnic	GWAS	D: 63,412 R: 176,890	Choquet et al., 2018 [111]
rs560713	*PDE7B*	phosphodiesterase 7B	604645	6q23.3	6	136,505,036	136,183,898		1				D: EU	GWAS	D: 6236	Ozel et al., 2014 [110]
rs190298731	*UST*	uronyl 2-sulfotransferase	610752	6q25.1	6	149,059,291	148,738,155	1					D: AA D: HI	GWAS	D: 10,961	Hoffmann et al., 2014 [172]
rs6917589	*SOD2*	superoxide dismutase 2, mitochondrial	147460	6q25.3	6	160,099,260	159,678,228	1					CH	AS	1413	Zhou et al., 2015 [354]
rs5746136	*SOD2*	superoxide dismutase 2, mitochondrial	147460	6q25.3	6	160,103,084	159,682,052	1					CH	AS	1413	Zhou et al., 2015 [354]
rs4880	*SOD2*	superoxide dismutase 2, mitochondrial	147460	6q25.3	6	160,113,872	159,692,840	1	1	1			SA	AS	629	Abu-Amero et al., 2014 [355]
		-		6q27	6	164,500,000 - 171,115,067	164,100,000 - 170,805,979		1				BDES	LS		Duggal et al., 2005 [188]
rs41269593	*RNASET2*	ribonuclease T2	612944	6q27	6	167,343,204	166,929,716		1				US-Cau	GWAS	D: 1660	Chen et al., 2015 [143]
indel	*THBS2*	thrombospondin II	188061	6q27	6	169,553,553	169,153,458				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs59072263	*GLCCI1/ICA1*	glucocorticoid-induced transcript 1 / islet cell autoantigen 1	614283 /147625	7p21.3	7	8,152,067	8,112,437		1				D: EU R: EU	GWAS	D: 2175 R: 4866	Blue Mountains Eye Study, 2013 [356]
rs12699251	*THSD7A*	thrombospondin type 1 domain containing 7A	612249	7p21.3	7	11,679,113	11,639,486	1					D: multi ethnic R: US, L, AS, AA	GWAS	D: 176,890 R: 63,412	Choquet et al., 2018 [111]
rs10274998	*DGKB*	diacylglycerol kinase beta	604070	7p21.2	7	14,245,377	14,205,752			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs11763147	*VKORC1L1*	vitmain K epoxide reductase complex, subunit 1-like 1	608838	7q11.21	7	65,326,821	65,861,834				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
indel	*RABGEF1*	rab guanine nucleotide exchange factor 1	609700	7q11.21	7	66,262,284	66,797,297				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	Locus/Gene/ Nearest gene (SNPs)	Gene name	MIM number	Locus	Chr	Hg19 (2009)	Hg38 (2013)	POAG	IOP	VCDR	CCT	ODA or OCA	Population	Study Type	Study size	Reference
rs4718428	*TMEM248*	transmembrane protein 248		7q11.21	7	66,421,446	66,956,459				1		D:AS R:AS	GWAS	D:7711 R:2681	Cornes et al., 2012 [316]
rs4718428	*TMEM248*	transmembrane protein 248		7q11.21	7	66,421,446	66,956,459				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
rs149154973	*ELN*	elastin	130160	7q11.23	7	73,322,211	73,907,882	1					D: US, L, AS, AA R: multi ethnic	GWAS	D: 63,412 R: 176,890	Choquet et al., 2018 [111]
rs1509922	*SEMA3C/HGF*	semaphorin 3C / hepatocyte growth factor	602645/142409	7q21.11	7	80,858,944	81,229,628		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs1045642	*ABCB1*	ATP binding cassette subfamily B member 1	171050	7q21.12	7	87,138,645	87,509,329	1	1				CH	AS	250	Liu, H. et al., 2016 [357]
rs2032582	*ABCB1*	ATP binding cassette subfamily B member 1	171050	7q21.12	7	87,160,618	87,531,302	1	1				CH	AS	250	Liu, H. et al., 2016 [357]
rs2106166	*SAMD9*	sterile alpha motif domain-containing protein 9	610456	7q21.2	7	92,668,332	93,039,018				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs662	*PON1*	paraoxonase 1	168820	7q21.3	7	94,937,446	95,308,134	N					JA	AS	839	Inagaki et al., 2006a [358]
rs6968419	*TFEC/TES*	transcription factor EC / Testis-derived transcript	604732/606085	7q31.2	7	115,823,384	116,183,330	1	1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs1052990	*CAV1/CAV2*	caveolin 1/caveolin 2	601047 /601048	7q31.2	7	116,148,370	116,508,316	1					D: IC R1:SW, UK, AU R2:CH	GWAS	D: 36140 R1: 4239 R2: 879	Thorleifsson et al., 2010 [77]
rs1052990	*CAV1/CAV2*	caveolin 1/caveolin 2	601047 /601048	7q31.2	7	116,148,370	116,508,316	1					US-Cau	AS	2183	Wiggs et al., 2011 [87]
rs10258482	*CAV1/CAV2*	caveolin 1/caveolin 2	601047 /601048	7q31.2	7	116,150,095	116,510,041	1	1				D: AS D: EU	GWAS	D: 7738 D:27558	Hysi et al., 2014 [86]
rs10262524	*CAV1/CAV2*	caveolin 1/caveolin 2	601047 /601048	7q31.2	7	116,150,952	116,510,898	1	1				D: AS D: EU	GWAS	D: 7738 D:27558	Hysi et al., 2014 [86]
rs10281637	*CAV1/CAV2*	caveolin 1/caveolin 2	601047 /601048	7q31.2	7	116,151,338	116,511,284	1	1				EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs17588172	*CAV1/CAV2*	caveolin 1/caveolin 2	601047 /601048	7q31.2	7	116,154,015	116,513,961	1					US-Cau	AS	6538	Loomis et al., 2014 [88]
rs17588172	*CAV1/CAV2*	caveolin 1/caveolin 2	601047 /601048	7q31.2	7	116,154,015	116,513,961	1	1				KO	AS	1161	Kim et al., 2015 [359]
rs6969706	*CAV1/CAV2*	caveolin 1/caveolin 2	601047 /601048	7q31.2	7	116,154,831	116,514,777	1					D: multi ethnic R: US, L, AS, AA	GWAS	D: 176,890 R: 63,412	Choquet et al., 2018 [111]
rs4236601	*CAV1/CAV2*	caveolin 1/caveolin 2	601047 /601048	7q31.2	7	116,162,729	116,522,675	1	1				D: IC R1:SW, UK, AU R2:CH	GWAS	D: 36140 R1: 4239 R2: 879	Thorleifsson et al., 2010 [77]
rs4236601	*CAV1/CAV2*	caveolin 1/caveolin 2	601047 /601048	7q31.2	7	116,162,729	116,522,675	0					US	AS	842	Kuehn et al., 2011 [360]
rs4236601	*CAV1/CAV2*	caveolin 1/caveolin 2	601047 /601048	7q31.2	7	116,162,729	116,522,675	1					US-Cau	AS	2183	Wiggs et al., 2011 [87]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	**Locus/Gene/ Nearest gene (SNPs)**	**Gene name**	**MIM number**	**Locus**	**Chr**	**Hg19 (2009)**	**Hg38 (2013)**	**POAG**	**IOP**	**VCDR**	**CCT**	**ODA or OCA**	**Population**	**Study Type**	**Study size**	**Reference**
rs4236601	*CAV1/CAV2*	caveolin 1/caveolin 2	601047 /601048	7q31.2	7	116,162,729	116,522,675	0					AF	AS	437	Cao et al., 2012 [96]
rs4236601	*CAV1/CAV2*	caveolin 1/caveolin 2	601047 /601048	7q31.2	7	116,162,729	116,522,675	0					SA	AS	625	Abu-Amero et al., 2012 [361]
rs4236601	*CAV1/CAV2*	caveolin 1/caveolin 2	601047 /601048	7q31.2	7	116,162,729	116,522,675	1					D: EU R: EU	GWAS	D: 3077 R: 13044	Gharahkhani et al., 2014 [89]
rs4236601	*CAV1/CAV2*	caveolin 1/caveolin 2	601047 /601048	7q31.2	7	116,162,729	116,522,675	1					US-Cau	AS	6538	Loomis et al., 2014 [88]
rs4236601	*CAV1/CAV2*	caveolin 1/caveolin 2	601047 /601048	7q31.2	7	116,162,729	116,522,675	0					BMES	AS	Cases: 67 C: 1919	Burdon et al 2015 [95]
rs4236601	*CAV1/CAV2*	caveolin 1/caveolin 2	601047 /601048	7q31.2	7	116,162,729	116,522,675	1					US, L, AS, AA	AS	63,412	Choquet et al., 2018
rs7795356	*CAV1/CAV2*	caveolin 1/caveolin 2	601047 /601048	7q31.2	7	116,217,029	116,576,975	1					D:JA R:JA	GWAS	D: 7993 R: 9014	Osman et al., 2012 [90]
rs76481776	*MIR182*	microRNA 182	611607	7q32.2	7	129,770,387	129,410,227	1	1				EU	AS	37,063	Liu, Y. et al., 2016 [288]
rs2070744	*NOS3*	nitric oxide synthase 3	163729	7q36	7	150,690,079	150,992,991	1					EG	AS	270	Emam et al., 2014 [362]
promoter	*NOS3*	nitric oxide synthase 3	163729	7q36	7	150,690,188 - 150,690,892	150,993,100 - 150,993,804	1					AU	AS	156	Tunny et al., 1998 [363]
	*ASB10 (*GLC1F*)*	ankyrin repeat- and SOCS box-containing protein 10	615054	7q35 – 7q36	7	150,872,785 - 150,884,919	151,175,698 - 151,187,832	1						LS		Locus: Wirtz et al., 1999 [71] Gene: Pasutto et al., 2012 [72]
rs3800791	*ASB10*	ankyrin repeat- and socs box-containing protein 10	615054	7q36.1	7	150,873,246	151,176,159	1					PA	AS	389	Micheal at al., 2015b [364]
rs2253592	*ASB10*	ankyrin repeat- and socs box-containing protein 10	615054	7q36.1	7	150,878,260	151,181,173	1					PA	AS	389	Micheal at al., 2015b [364]
rs3808520	*LOXL2*	lysyl oxidase-like 2	606663	8p21.3	8	23,164,773	23,307,260				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs11759831	*CRISPLD1*	cysteine-rich secretory protein LCCL domain containing 1	-	8q21.11	8	48,896,220	48,928,583			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
indel	*PKIA*	protein kinase (cAMP-dependent, catalytic) inhibitor alpha	606059	8q21.13	8	78,380,944	77,468,708		1				EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs7815043	*PKIA*	protein kinase (cAMP-dependent, catalytic) inhibitor alpha	606059	8q21.13	8	78,932,025	78,019,790		1				EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs6468996	*DCAF4L2*	DDB1 and CUL4 associated factor 4-like 2	-	8q21.3	8	88,735,337	87,723,109			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
indel	*DCAF4L2*	DDB1 and CUL4 associated factor 4-like 2	-	8q21.3	8	88,744,441	87,732,213			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs9969524	*DCAF4L2*	DDB1 and CUL4 associated factor 4-like 2	-	8q21.3	8	88,746,846	87,734,618					1	EU AS	GWAS	17248 6841	Springelkamp et al., 2015 [134]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	**Locus/Gene/ Nearest gene (SNPs)**	**Gene name**	**MIM number**	**Locus**	**Chr**	**Hg19 (2009)**	**Hg38 (2013)**	**POAG**	**IOP**	**VCDR**	**CCT**	**ODA or OCA**	**Population**	**Study Type**	**Study size**	**Reference**
rs10429294	*NDUFAF6*	NADH:ubiquinone oxidoreductase complex assembly factor 6	612392	8q22.1	8	95,969,322	94,957,094				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs284489	*ZFPM2/LRP12*	zinc finger protein multitype 2/ low density lipoprotein receptor-related protein 12	603693/-	8q22.3	8	105,958,020	104,945,792	1 N					D: EU	GWAS	D: 6633	Wiggs et al., 2012 [94]
	GLC1D		602429	8q23	8	106,200,001 - 117,700,000	105,100,001 - 116,700,000	1					US	LS	1 large family	Trifan et al., 1998 [173]
rs2514884	*ANGPT1*	Angiopoietin 1	601667	8q23.1	8	108,276,873	107,264,645	1					D: multi ethnic R: US, L, AS, AA	GWAS	D: 176,890 R: 63,412	Choquet et al., 2018 [111]
rs10105844	*ANGPT1*	Angiopoietin 1	601667	8q23.1	8	108,290,872	107,278,644		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs7815720	*SLC30A8*	solute carrier family 30 (zinc transporter), member 8	611145	8q24.11	8	118,089,513	117,077,274		1				D: EU	GWAS	D: 6236	Ozel et al., 2014 [110]
rs61753798	*MTBP*	MDM2 binding protein	605927	8q24.12	8	121,530,092	120,517,852		1				US-Cau	GWAS	D: 1660	Chen et al., 2015 [143]
rs2945733	*ST3GAL1*	ST3 Beta-galactoside alpha-2,3-sialyltransferase 1	607187	8q24.22	8	134,615,750	133,603,507		1				D: EU	GWAS	D: 6236	Ozel et al., 2014 [110]
rs2920293	*PSCA*	prostate stem cell antigen	602470	8q24.3	8	143,765,414	142,683,996			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs189033490	*FAM83H*	family with sequence similarity 83, member H	611927	8q24.3	8	144,811,340	143,729,170		1				US-Cau	GWAS	D: 1660	Chen et al., 2015 [143]
rs7026684	*GLIS3*	GLIS family zinc finger 3	610192	9p24.2	9	4,215,308	4,215,308				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs2224492	*GLIS3*	GLIS family zinc finger protein 3	610192	9p24.2	9	4,237,546	4,237,546		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs1324183	*MPDZ/NFIB*	multiple PDZ domain protein/nuclear factor 1/B	603785/?	9p23	9	13,557,491	13,557,492				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
rs1831902	*LINC01235/ LINC00583*	long intergenic non-protein coding RNA 1235 / long intergenic non-protein coding RNA 583		9p23	9	13,558,317	13,558,318		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs66720556	*MPDZ*	Multiple PDZ domain protein	603785	9p23	9	13,559,717	13,559,718				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs1063192	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,003,367	22,003,368			1			D: NL R: NL, UK	GWAS	D: 7360 R: 4455	Ramdas et al., 2010 [99]
rs1063192	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,003,367	22,003,368	1		1			US-Cau	AS	875	Fan et al., 2011 [97]
rs1063192	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,003,367	22,003,368	1					EU	AS	45998	Ramdas et al., 2011 [69]
rs1063192	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,003,367	22,003,368	1					AF	AS	437	Cao et al., 2012 [96]
rs1063192	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,003,367	22,003,368	1		1			AU, NZ	AS	1759	Dimasi et al., 2012 [29]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	Locus/Gene/ Nearest gene (SNPs)	Gene name	MIM number	Locus	Chr	Hg19 (2009)	Hg38 (2013)	POAG	IOP	VCDR	CCT	ODA or OCA	Population	Study Type	Study size	Reference
rs1063192	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,003,367	22,003,368	N		1			JA	AS	616	Mabuchi et al., 2012 [98]
rs1063192	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,003,367	22,003,368	1					D:JA R:JA	GWAS	D: 7993 R: 9014	Osman et al., 2012 [90]
rs1063192	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,003,367	22,003,368	1					BMES	AS	OAG: 67 C: 1919	Burdon et al 2015 [95]
rs523096	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,019,129	22,019,130	N					D:JA R:JA	GWAS	D: 1519 R: 700	Nakano et al., 2012 [101]
rs523096	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,019,129	22,019,130	N					D:JA R:JA	GWAS	D: 843 R: 697	Takamoto et al., 2012 [102]
rs7049105	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,028,801	22,028,802	N	1	1			AU, NZ	AS	2027	Burdon et al., 2012 [92]
rs7049105	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,028,801	22,028,802	N					US-Cau	GWAS	6633	Wiggs et al., 2012 [94]
rs7865618	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,031,005	22,031,006	1					D: JA R: JA	GWAS	D: 1519 R: 700	Nakano et al., 2012 [101]
rs7865618	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,031,005	22,031,006			1			EU AS	GWAS	27,878	Springelkamp et al., 2014 [103]
rs2157719	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,033,366	22,033,367	N					US-Cau	GWAS	6633	Wiggs et al., 2012 [94]
rs2157719	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,033,366	22,033,367	1					AS, AF, EU	GWAS	D: 13250 R: 35953	Li et al., 2015 [100]
rs2157719	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,033,366	22,033,367	1		1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs1333037	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,040,765	22,040,766	1					D: multi ethnic R: US, L, AS, AA	GWAS	D: 176,890 R: 63,412	Choquet et al., 2018 [111]
rs1412829	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,043,926	22,043,927	1					D:AU,NZ R:AU,NZ	GWAS	D: 4546 R: 4148	Burdon et al., 2011 [91]
rs1412829	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,043,926	22,043,927	1					BMES	AS	OAG: 67 C: 1919	Burdon et al 2015 [95]
rs1360589	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,045,317	22,045,318	1		1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs10811645	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,049,656	22,049,657	1 N					D: US, L, AS, AA R: multi ethnic	GWAS	D: 63,412 R: 176,890	Choquet et al., 2018 [111]
rs944800	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,050,898	22,050,899	1					Japan	GWAS	POAG:7378 C:36385	Shiga et al., 2018 [337]
rs7866783	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,056,359	22,056,360	1					Meta-analysis US, AU, EU, SC	GWAS	37333	Bailey et al., 2016 [12]
rs10120688	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,056,499	22,056,500	1	1	1			D:AU,NZ R:AU,NZ	GWAS	D: 4546 R: 4148	Burdon et al., 2011 [91]
rs10120688	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,056,499	22,056,500	N	1	1			AU, NZ	AS	2027	Burdon et al., 2012 [92]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	Locus/Gene/ Nearest gene (SNPs)	Gene name	MIM number	Locus	Chr	Hg19 (2009)	Hg38 (2013)	POAG	IOP	VCDR	CCT	ODA or OCA	Population	Study Type	Study size	Reference
rs10120688	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,056,499	22,056,500	1					D:UK R:UK	GWAS	D: 437 R: 344	Gibson et al., 2012 [93]
rs10120688	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,056,499	22,056,500	N					US-Cau	GWAS	6633	Wiggs et al., 2012 [94]
rs4977756	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,068,652	22,068,653	1	1	1			D:AU,NZ R:AU,NZ	GWAS	D: 4546 R: 4148	Burdon et al., 2011 [91]
rs4977756	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,068,652	22,068,653	0					AF	AS	437	Cao et al., 2012 [96]
rs4977756	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,068,652	22,068,653	N					US-Cau	GWAS	6633	Wiggs et al., 2012 [94]
rs4977756	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,068,652	22,068,653	1					D: EU R: EU	GWAS	D: 3077 R: 13044	Gharahkhani et al., 2014 [89]
rs4977756	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,068,652	22,068,653	1					BMES	AS	OAG: 67 C: 1919	Burdon et al 2015 [95]
rs4977756	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,068,652	22,068,653	1					JA	AS	523	Kimura et al., 2015 [365]
rs4977756	*CDKN2B-AS1*	CDKN2B antisense RNA 1	613149	9p21.3	9	22,068,652	22,068,653	1					US, L, AS, AA	AS	63,412	Choquet et al., 2018 [111]
	GLC1J		608695	9q22	9	90,400,001 - 102,600,000	87,800,001 - 99,800,000	J					US	LS	several families	Wiggs et al., 2004 [47]
rs2487048	*ABCA1*	ATP-Binding cassette, subfamily A, member 1	600046	9q31.1	9	107,691,823	104,929,542	1					EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs2472496	*ABCA1*	ATP-Binding cassette, subfamily A, member 1	600046	9q31.1	9	107,695,353	104,933,072		1	1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs2472494	*ABCA1*	ATP-Binding cassette, subfamily A, member 1	600046	9q31.1	9	107,695,539	104,933,258	1					Japan	GWAS	POAG:7378 C:36385	Shiga et al., 2018 [337]
rs2472493	*ABCA1*	ATP-Binding cassette, subfamily A, member 1	600046	9q31.1	9	107,695,848	104,933,567	1					D: EU R: EU	GWAS	D: 3077 R: 13044	Gharahkhani et al., 2014 [89]
rs2472493	*ABCA1*	ATP-Binding cassette, subfamily A, member 1	600046	9q31.1	9	107,695,848	104,933,567		1				D: AS D: EU	GWAS	D: 7738 D:27558	Hysi et al., 2014 [86]
rs2472493	*ABCA1*	ATP-Binding cassette, subfamily A, member 1	600046	9q31.1	9	107,695,848	104,933,567	1					Meta-analysis US, AU, EU, SC	GWAS	37333	Bailey et al., 2016 [12]
rs2472493	*ABCA1*	ATP binding cassette subfamily A member 1	600046	9q31.1	9	107,695,848	104,933,567	1	1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs2472493	*ABCA1*	ATP-Binding cassette, subfamily A, member 1	600046	9q31.1	9	107,695,848	104,933,567	1	1				D: US, L, AS, AA R: multi ethnic	GWAS	D: 63,412 R: 176,890	Choquet et al., 2018 [111]
rs2472493	*ABCA1*	ATP-Binding cassette, subfamily A, member 1	600046	9q31.1	9	107,695,848	104,933,567	1					D: multi ethnic R: US, L, AS, AA	GWAS	D: 176,890 R: 63,412	Choquet et al., 2018 [111]
rs2472493	*ABCA1*	ATP-Binding cassette, subfamily A, member 1	600046	9q31.1	9	107,695,848	104,933,567	1					US, L, AS, AA	AS	63,412	Choquet et al., 2018 [111]
rs2487032	*ABCA1*	ATP-Binding cassette, subfamily A, member 1	600046	9q31.1	9	107,703,934	104,941,653		1				D: CH R: CH	GWAS	D: 1971 R: 3337	Chen et al., 2014 [366]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	Locus/Gene/ Nearest gene (SNPs)	Gene name	MIM number	Locus	Chr	Hg19 (2009)	Hg38 (2013)	POAG	IOP	VCDR	CCT	ODA or OCA	Population	Study Type	Study size	Reference
rs10980623	*LPAR1*	lysophosphatidic acid receptor 1	602282	9q31.3	9	113,660,537	110,898,257				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs1007000	*LPAR1*	lysophosphatidic acid receptor 1	602282	9q31.3	9	113,662,681	110,900,401				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
rs2149356	*TLR4*	toll-like receptor 4	603030	9q33.1	9	120,474,199	117,711,921	1					JA	AS	765	Takano et al., 2012 [367]
rs7037117	*TLR4*	toll-like receptor 4	603030	9q33.1	9	120,483,663	117,721,385	N					JA	AS	568	Shibuya et al., 2008 [368]
rs7037117	*TLR4*	toll-like receptor 4	603030	9q33.1	9	120,483,663	117,721,385	0					AF	AS	437	Cao et al., 2012 [96]
rs7037117	*TLR4*	toll-like receptor 4	603030	9q33.1	9	120,483,663	117,721,385	1					CH	AS	1039	Chen et al., 2012b [326]
rs7037117	*TLR4*	toll-like receptor 4	603030	9q33.1	9	120,483,663	117,721,385	1					JA	AS	765	Takano et al., 2012 [367]
rs2286885	*MVB12B (FAM125B)*	multivesicular body subunit 12B		9q33.3	9	129,246,487	126,484,208		1				UK	AS	2774	Nag et al., 2014 [129]
rs6478746	*LMX1B*	LIM homeobox transcription factor 1, beta	602575	9q33.3	9	129,367,398	126,605,119	1				1	EU	GWAS	OAG:3071 C:6750	Gharahkhani et al., 2018 [131]
rs10819187	*LMX1B*	LIM homeobox transcription factor 1, beta	602575	9q34.1	9	129,369,971	126,607,692	1					Japan	GWAS	POAG:7378 C:36385	Shiga et al., 2018 [337]
rs55770306	*LMX1B*	LIM homeobox transcription factor 1, beta	602575	9q34.1	9	129,388,033	126,625,754	1					D: multi ethnic R: US, L, AS, AA	GWAS	D: 176,890 R: 63,412	Choquet et al., 2018 [111]
rs7854658	*LMX1B*	LIM homeobox transcription factor 1, beta	602575	9q34.1	9	129,414,938	126,652,659	1					UK	AS	643	Park et al., 2009 [369]
rs8176743	*ABO*	ABO glycosyltransferase	110300	9q34.2	9	136,131,415	133,256,028		1				D: AS D: EU	GWAS	D: 7738 D:27558	Hysi et al., 2014 [86]
rs8176741	*ABO*	ABO glycosyltransferase	110300	9q34.2	9	136,131,461	133,256,074		1	1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs3094339	*VAV2/BRD3*	vav 2 guanine nucleotide exchange factor / bromodomain containing protein 3	600428 / 601541	9q34.2	9	136,884,738	134,019,616				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs4841899	*RXRA/COL5A1*	retinoid X receptor alpha / collagen, type V, alpha-1	180245 / 120215	9q34.2/ 9q34.3	9	137,424,412	134,532,566				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs3132306	*RXRA/COL5A1*	retinoid X receptor alpha / collagen, type V, alpha-1	180245 / 120215	9q34.2/ 9q34.3	9	137,440,212	134,548,366				1		D: AS R: AS	GWAS	D: 7711 R: 2681	Cornes et al., 2012 [316]
rs3132306	*RXRA/COL5A1*	retinoid X receptor alpha / collagen, type V, alpha-1	180245 / 120215	9q34.2/ 9q34.3	9	137,440,212	134,548,366				1		D: GE R: NL	GWAS	D: 3931 R: 1418	Hoehn et al., 2012 [198]
rs1536482	*RXRA/COL5A1*	retinoid X receptor alpha / collagen, type V, alpha-1	180245 / 120215	9q34.2/ 9q34.3	9	137,440,528	134,548,682				1		D: CR, SCL R: CR, SCL	GWAS	D: 1445 R: 824	Vitart et al., 2010 [201]
rs1536482	*RXRA/COL5A1*	retinoid X receptor alpha / collagen, type V, alpha-1	180245 / 120215	9q34.2/ 9q34.3	9	137,440,528	134,548,682				1		D1: SM D2: SG-CH	GWAS	D1: 3280 D2: 3400	Vithana et al., 2011 [315]
rs1536482	*RXRA/COL5A1*	retinoid X receptor alpha / collagen, type V, alpha-1	180245 / 120215	9q34.2/ 9q34.3	9	137,440,528	134,548,682				1		D: AS R: AS	GWAS	D: 7711 R: 2681	Cornes et al., 2012 [316]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	**Locus/Gene/ Nearest gene (SNPs)**	**Gene name**	**MIM number**	**Locus**	**Chr**	**Hg19 (2009)**	**Hg38 (2013)**	**POAG**	**IOP**	**VCDR**	**CCT**	**ODA or OCA**	**Population**	**Study Type**	**Study size**	**Reference**
rs1536482	*RXRA/COL5A1*	retinoid X receptor alpha / collagen, type V, alpha-1	180245 / 120215	9q34.2/ 9q34.3	9	137,440,528	134,548,682				0		AU, NZ	AS	1759	Dimasi et al., 2012 [29]
rs1536482	*RXRA/COL5A1*	retinoid X receptor alpha / collagen, type V, alpha-1	180245 / 120215	9q34.2/ 9q34.3	9	137,440,528	134,548,682				1		D: GE R: NL	GWAS	D: 3931 R: 1418	Hoehn et al., 2012 [198]
rs1536482	*RXRA/COL5A1*	retinoid X receptor alpha / collagen, type V, alpha-1	180245 / 120215	9q34.2/ 9q34.3	9	137,440,528	134,548,682	0			0		D: US-Cau AS: US-Cau	GWAS AS	D: 1117 AS: 6469	Ulmer et al., 2012 [196]
rs1536482	*RXRA/COL5A1*	retinoid X receptor alpha / collagen, type V, alpha-1	180245 / 120215	9q34.2/ 9q34.3	9	137,440,528	134,548,682				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
rs1536482	*RXRA/COL5A1*	retinoid X receptor alpha / collagen, type V, alpha-1	180245 / 120215	9q34.2/ 9q34.3	9	137,440,528	134,548,682				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs3118520	*RXRA/COL5A1*	retinoid X receptor alpha / collagen, type V, alpha-1	180245 / 120215	9q34.2/ 9q34.3	9	137,440,528	134,548,682				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
rs3132303	*RXRA/COL5A1*	retinoid X receptor alpha / collagen, type V, alpha-1	180245 / 120215	9q34.2/ 9q34.3	9	137,444,298	134,552,452				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs7032489	*COL5A1*	collagen, type V, alpha-1	120215	9q34.3	9	137,559,775	134,667,929				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs7044529	*COL5A1*	collagen, type V, alpha-1	120215	9q34.3	9	137,568,051	134,676,205				1		D1: SM D2: SG-CH	GWAS	D1: 3280 D2: 3400	Vithana et al., 2011 [315]
rs7044529	*COL5A1*	collagen, type V, alpha-1	120215	9q34.3	9	137,568,051	134,676,205				0		AU, NZ	AS	1759	Dimasi et al., 2012 [29]
rs7044529	*COL5A1*	collagen, type V, alpha-1	120215	9q34.3	9	137,568,051	134,676,205	0			0		D: US-Cau AS: US-Cau	GWAS AS	D: 1117 AS: 6469	Ulmer et al., 2012 [196]
rs7044529	*COL5A1*	collagen, type V, alpha-1	120215	9q34.3	9	137,568,051	134,676,205				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
rs7040970	*LCN12/PTGDS*	Lipoclalin 12/Prostaglandin D2 synthase, brain	612905 / 176803	9q34.3	9	139,859,013	136,964,561				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs11145951	*LCN12/PTGDS*	Lipoclalin 12/Prostaglandin D2 synthase, brain	612905 / 176803	9q34.3	9	139,860,264	136,965,812				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
	*OPTN (*GLC1E*)*	Optineurin	602432	10p13	10	13,142,082-13,180,276	13,100,082-13,138,276	N					UK	LS	1 large family	Locus: Sarfarazi et al., 1998 [70] Gene: Rezaie et al., 2002 [59]
	–			10p12.33–10p12.1	10	17,300,001-29,600,000	17,300,001-29,300,000	1					AC	LS	several families	Nemesure et al., 2003 [187]
rs7098387	*PLXDC2*	plexin domain containing 2	606827	10p12.31	10	20,623,473	20,334,544	1					KO	AS	1115	Kim et al., 2014 [327]
rs7081455	*PLXDC2*	plexin domain containing 2	606827	10p12.31	10	20,638,885	20,349,956	0					AF	AS	437	Cao et al., 2012 [96]
rs7081455	*PLXDC2*	plexin domain containing 2	606827	10p12.31	10	20,638,885	20,349,956	0					CH	AS	1039	Chen et al., 2012b [326]
rs7081455	*PLXDC2*	plexin domain containing 2	606827	10p12.31	10	20,638,885	20,349,956	1					D: JA R: JA	GWAS	D: 1519 R: 857	Nakano et al., 2012 [101]
rs11014632	*GPR158*	g protein-coupled receptor 158	614573	10p12.1	10	25,877,651	25,588,722		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	Locus/Gene/ Nearest gene (SNPs)	Gene name	MIM number	Locus	Chr	Hg19 (2009)	Hg38 (2013)	POAG	IOP	VCDR	CCT	ODA or OCA	Population	Study Type	Study size	Reference
rs7916852	*MPP7*	membrane protein, palmitoylated 7	610973	10p12.1	10	28,405,411	28,116,482	1					IN	GWAS	D: 729 R: 714	Vishal et al., 2016 [370]
rs7090871	*ARID5B*	AT rich interactive domain 5B	608538	10q21.2	10	63,830,286	62,070,527				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
rs35809595	*ARID5B*	AT-rich interactive domain 5B	608538	10q21.2	10	63,831,928	62,072,169				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs5785510*	*ARID5B*	AT-rich interaction domain-containing protein 5B	608538	10q21.2	10	63,837,290	62,077,531		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs10761970	*CTNNA3*	catenin alpha 3	607667	10q21.3	10	67,160,860	65,401,102		1				NIES	GWAS (Ped.)	330	Matovinovic et al., 2017 [133]
rs61854782	*ATOH7*	atonal homolog 7	609875	10q21.3	10	69,991,749	68,231,992			1			CH	AS	431	Chen et al., 2012a [371]
rs7916697	*ATOH7*	atonal homolog 7	609875	10q21.3	10	69,991,853	68,232,096	0					US-Cau	AS	875	Fan et al., 2011 [97]
rs7916697	*ATOH7*	atonal homolog 7	609875	10q21.3	10	69,991,853	68,232,096					1	D:AS R:NL	GWAS	D: 4445 R: 9326	Khor et al., 2011 [317]
rs7916697	*ATOH7*	atonal homolog 7	609875	10q21.3	10	69,991,853	68,232,096	1					AF	AS	437	Cao et al., 2012 [96]
rs7916697	*ATOH7*	atonal homolog 7	609875	10q21.3	10	69,991,853	68,232,096	0					CH	AS	431	Chen et al., 2012a [371]
rs7916697	*ATOH7*	atonal homolog 7	609875	10q21.3	10	69,991,853	68,232,096			1		1	AU, UK	AS	3599	Venturini et al., 2014 [130]
rs7916697	*ATOH7*	atonal homolog 7	609875	10q21.3	10	69,991,853	68,232,096			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs7916410	*ATOH7*	atonal homolog 7	609875	10q21.3	10	69,995,667	68,235,910			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs1900005	*ATOH7*	atonal homolog 7	609875	10q21.3	10	69,998,055	68,238,298			1			EU, AS	GWAS	27,878	Springelkamp et al., 2014 [103]
rs1900004	*ATOH7*	atonal homolog 7	609875	10q21.3	10	70,000,881	68,241,124			1		1	D:AU R:UK	GWAS	D: 1368 R: 848	Macgregor et al., 2010 [335]
rs1900004	*ATOH7*	atonal homolog 7	609875	10q21.3	10	70,000,881	68,241,124		1	1			IN	AS	468	Philomenadin et al., 2015 [105]
rs1900004	*ATOH7*	atonal homolog 7	609875	10q21.3	10	70,000,881	68,241,124					1	D:NL	GWAS	D: 23000	Axenovich et al., 2011 [104]
rs1900004	*ATOH7*	atonal homolog 7	609875	10q21.3	10	70,000,881	68,241,124	0					African	AS	437	Cao et al., 2012 [96]
rs1900004	*ATOH7*	atonal homolog 7	609875	10q21.3	10	70,000,881	68,241,124	0					CH	AS	431	Chen et al., 2012a [371]
rs1900004	*ATOH7*	atonal homolog 7	609875	10q21.3	10	70,000,881	68,241,124					0	AU, NZ	AS	1759	Dimasi et al., 2012 [29]
rs1900004	*ATOH7*	atonal homolog 7	609875	10q21.3	10	70,000,881	68,241,124					1	US-Cau	AS	875	Fan et al., 2011 [97]
rs1900004	*ATOH7*	atonal homolog 7	609875	10q21.3	10	70,000,881	68,241,124	N					JA	AS	616	Mabuchi et al., 2012 [98]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	**Locus/Gene/ Nearest gene (SNPs)**	**Gene name**	**MIM number**	**Locus**	**Chr**	**Hg19 (2009)**	**Hg38 (2013)**	**POAG**	**IOP**	**VCDR**	**CCT**	**ODA or OCA**	**Population**	**Study Type**	**Study size**	**Reference**
rs1900004	*ATOH7*	atonal homolog 7	609875	10q21.3	10	70,000,881	68,241,124			1		1	D:NL R:NL, UK	GWAS	D: 7360 R: 4455	Ramdas et al., 2010 [99]
rs1900004	*ATOH7*	atonal homolog 7	609875	10q21.3	10	70,000,881	68,241,124	1					EU	AS	45998	Ramdas et al., 2011 [69]
rs1900004	*ATOH7*	atonal homolog 7	609875	10q21.3	10	70,000,881	68,241,124					1	EU AS	GWAS	17248 6841	Springelkamp et al., 2015 [134]
rs3858145	*ATOH7*	atonal homolog 7	609875	10q21.3	10	70,011,838	68,252,081					1	D:AU R:UK	GWAS	D: 1368 R: 848	Macgregor et al., 2010 [335]
rs3858145	*ATOH7*	atonal homolog 7	609875	10q21.3	10	70,011,838	68,252,081			0			US-Cau	AS	875	Fan et al., 2011 [97]
rs3858145	*ATOH7*	atonal homolog 7	609875	10q21.3	10	70,011,838	68,252,081	0					AF	AS	437	Cao et al., 2012 [96]
rs3858145	*ATOH7*	atonal homolog 7	609875	10q21.3	10	70,011,838	68,252,081			1			CH	AS	431	Chen et al., 2012a [371]
rs3858145	*ATOH7*	atonal homolog 7	609875	10q21.3	10	70,011,838	68,252,081	0					AU, NZ	AS	1759	Dimasi et al., 2012 [29]
rs3858145	*ATOH7*	atonal homolog 7	609875	10q21.3	10	70,011,838	68,252,081			0			D:UK R:UK	GWAS	D: 437 R: 344	Gibson et al., 2012 [93]
rs3858145	*ATOH7*	atonal homolog 7	609875	10q21.3	10	70,011,838	68,252,081			1			EU AS	GWAS	17248 6841	Springelkamp et al., 2015 [134]
rs12571093	*ATOH7/PBLD*	atonal homolog 7/phenazine biosynthesis-like protein domain containing	609875 / 612189	10q21.3	10	70,019,371	68,259,614					1	D:AU R:UK	GWAS	D: 1368 R: 848	Macgregor et al. 2010 [335]
	–			10q22	10	70,600,001 - 82,000,000	68,800,001 - 80,300,000		1				AU	LS	1 large family	Charlesworth et al., 2005 [189]
rs72815193	*CYP26A1/ MYOF*	cytochrome P450, family 26, subfamily A, polypeptide 1 / myoferlin	602239 / 604603	10q23.33	10	94,963,391	93,203,634	1		1			EU	GWAS	OAG:3071 C:6750	Gharahkhani et al., 2018 [131]
rs66479974*	*CYP26A1/ MYOF*	cytochrome P450, family 26, subfamily A, polypeptide 1 / myoferlin	602239 / 604603	10q23.33	10	95,049,398	93,289,641		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
indel	*PLCE1*	phospholipase C, epsilon 1	608414	10q23.33	10	96,008,348	94,248,591			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs3891783	*PLCE1*	phospholipase C, epsilon 1	608414	10q23.33	10	96,015,793	94,256,036			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs1830890	*PLCE1*	phospholipase C, epsilon 1	608414	10q23.33	10	96,019,501	94,259,744			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs7072574	*PLCE1*	phospholipase C, epsilon 1	608414	10q23.33	10	96,036,306	94,276,549			1			EU, AS	GWAS	27,878	Springelkamp et al., 2014 [103]
rs2274224	*PLCE1*	phospholipase C, epsilon 1	608414	10q23.33	10	96,039,597	94,279,840	1					D: multi ethnic R: US, L, AS, AA	GWAS	D: 176,890 R: 63,412	Choquet et al., 2018
rs2419835	*HABP2*	hyaluronan binding protein 2	603924	10q25.3	10	115,296,564	113,536,805				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	Locus/Gene/ Nearest gene (SNPs)	Gene name	MIM number	Locus	Chr	Hg19 (2009)	Hg38 (2013)	POAG	IOP	VCDR	CCT	ODA or OCA	Population	Study Type	Study size	Reference
rs1801253	*ADRB1*	beta-1-adrenergic receptor	109630	10q24–10q26	10	115,805,056	114,045,297	N					JA	AS	745	Inagaki et al., 2006b [186]
rs1681739	*HSPA12A*	heat shock 70kDa protein 12A	610701	10q25.3	10	118,563,329	116,803,818			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs12262706	*LHPP*	phospholysine phosphohistidine inorganic pyrophosphate phosphatase	617231	10q26.13	10	126,247,468	124,558,899	1					Japan	GWAS	POAG:7378 C:36385	Shiga et al., 2018 [337]
exon 9	*IGF2*	insulin-like growth factor II	147470	11p15.5	11	2,154,218 - 2,154,895	2,132,987 - 2,133,667	1					CH	AS	164	Tsai et al., 2003 [372]
rs2030324	*BDNF*	brain-derived neurotrophic factor	113505	11p14.1	11	27,726,915	27,705,368	1		1			PO	AS	769	Nowak et al., 2015 [351]
rs542340	*DNAJC24*	DnaJ/Hsp40 homolog, subfamily C, member 24	611072	11p13	11	31,409,438	31,387,891	1					D: AA D: HI	GWAS	D: 10,961	Hoffmann et al., 2014 [172]
rs542340	*DNAJC24*	DnaJ/Hsp40 homolog, subfamily C, member 24	611072	11p13	11	31,409,438	31,387,891	0					US, L, AS, AA	AS	63,412	Choquet et al., 2018 [111]
rs1223068	*IMMP1L*	inner mitochondrial membrane peptidase subunit 1	612323	11p13	11	31,480,349	31,458,802			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs555091	*ELP4*	elongator acetyltransferase complex, subunit 4	606985	11p13	11	31,532,959	31,511,412	1					D: AA D: HI	GWAS	D: 10,961	Hoffmann et al., 2014 [172]
rs11031436	*ELP4*	elongator acetyltransferase complex, subunit 4	606985	11p13	11	31,663,882	31,642,334					1	EU AS	GWAS	17248 6841	Springelkamp et al., 2015 [134]
rs7126851	*PAX6*	paired box gene 6	607108	11p13	11	31,753,946	31,732,398					1	D: NL R: NL	GWAS	D: 5312 R: 5933	Gasten et al., 2012 [373]
rs7104512	*PAX6*	paired box gene 6	607108	11p13	11	31,768,884	31,747,336					1	D: NL R: NL	GWAS	D: 5312 R: 5933	Gasten et al., 2012 [373]
rs10835818	*PAX6*	paired box gene 6	607108	11p13	11	31,793,051	31,771,503					1	D: NL R: NL	GWAS	D: 5312 R: 5933	Gasten et al., 2012 [373]
rs3026398	*PAX6*	paired box gene 6	607108	11p13	11	31,808,775	31,787,227				1		AU-Cau	AS	956	Dimasi et al., 2010 [374]
rs1001179	*CAT*	catalase	115500	11p13	11	34,460,231	34,438,684	1					SA	AS	628	Abu-Amero et al., 2013 [375]
rs79390637	*PSMC3/RAPSN*	proteasome 26S subunit, ATPase, 3 / receptor-associated protein of the synapse	186852 / 601592	11p11.2	11	47,456,867	47,435,316		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs12419342	*RAPSN*	receptor-associated protein of the synapse	601592	11p11.2	11	47,468,545	47,446,993		1				D: AS D: EU	GWAS	D: 7738 D:27558	Hysi et al., 2014 [86]
rs747782	*NUP160 / PTPRJ*	nucleoporin, 160-KD / protein-tyrosine phostphatase, receptor-type, J	607614 / 600925	11p11.2	11	47,940,925	47,919,373		1				D: AS D: EU	GWAS	D: 7738 D:27558	Hysi et al., 2014 [86]
rs7944735	*NUP160 / PTPRJ*	nucleoporin, 160-KD / protein-tyrosine phostphatase, receptor-type, J	607614 / 600925	11p11.2	11	47,955,608	47,934,056	1	1				EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	**Locus/Gene/ Nearest gene (SNPs)**	**Gene name**	**MIM number**	**Locus**	**Chr**	**Hg19 (2009)**	**Hg38 (2013)**	**POAG**	**IOP**	**VCDR**	**CCT**	**ODA or OCA**	**Population**	**Study Type**	**Study size**	**Reference**
rs1681630	*PTPRJ*	Protein-tyrosine phosphatase, receptor type J	600925	11p11.2	11	47,969,152	47,947,600		1				D: AS D: EU	GWAS	D: 7738 D:27558	Hysi et al., 2014 [86]
rs7946766	*PTPRJ*	Protein-tyrosine phosphatase, receptor type J	600925	11p11.2	11	48,004,369	47,982,817		1				D: AS D: EU	GWAS	D: 7738 D:27558	Hysi et al., 2014 [86]
rs17146964	*SCYL1*	SCY1-like 1	607982	11q13	11	65,249,145	65,481,674			1			D:NL R:NL, UK	GWAS	D: 7360 R: 4455	Ramdas et al. 2010 [99]
rs1346	*SSSCA1*	Sjogren syndrome/scleroderma autoantigen 1	606044	11q13.1	11	65,337,251	65,569,780			1			EU, AS	GWAS	27,878	Springelkamp et al., 2014 [103]
rs1346	*SSSCA1*	Sjogren syndrome/scleroderma autoantigen 1	606044	11q13.1	11	65,337,251	65,569,780	1		1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs790357	*DLG2*	discs, large homolog 2	603583	11q14.1	11	83,620,940	83,909,897		1				NIES	GWAS (Ped.)	330	Matovinovic et al., 2017 [133]
rs1799750	*MMP1*	matrix metalloproteinase 1	120353	11q22.2	11	102,670,496	102,799,765	1					PO	AS	449	Majsterek et al., 2011 [376]
rs4938174	*ARHGAP20 / C11orf53*	Rho GTPase activating protein 20 / chromosome 11 open reading frame 53	609568/-	11q23.1	11	110,913,240	111,042,516				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
rs4938174	*ARHGAP20 / C11orf53*	Rho GTPase activating protein 20 / chromosome 11 open reading frame 53	609568/-	11q23.1	11	110,913,240	111,042,516				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs12806740	*TMEM136*	transmembrane protein 136	-	11q23.3	11	120,203,628	120,332,919	1					D: multi ethnic R: US, L, AS, AA	GWAS	D: 176,890 R: 63,412	Choquet et al., 2018 [111]
rs58073046	*ARHGEF12*	Rho guanine nucleotide exchange factor 12	604763	11q23.3	11	120,248,493	120,377,784		1				D: EU	GWAS	D: 8015 R: 7471	Springelkamp et al., 2015 [106]
rs58073046	*ARHGEF12*	Rho guanine nucleotide exchange factor 12	604763	11q23.3	11	120,248,493	120,377,784	1					US, L, AS, AA	AS	63,412	Choquet et al., 2018 [111]
rs12794618	*ARHGEF12*	Rho guanine nucleotide exchange factor 12	604763	11q23.3	11	120,289,699	120,418,990		1				EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs2276035	*ARHGEF12*	Rho guanine nucleotide exchange factor 12	604763	11q23.3	11	120,346,360	120,475,651	1					D: EU R: EU	GWAS	D: 3077 R: 13044	Gharahkhani et al., 2014 [89]
rs199800298*	*ARHGEF12*	Rho guanine nucleotide exchange factor 12	604763	11q23.3	11	120,348,584	120,477,875		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
indel	*ARHGEF12*	Rho guanine nucleotide exchange factor 12	604763	11q23.3	11	120,357,425	120,486,716	1	1				EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs4936099	*ADAMTS8*	a disintegrin-like and metalloproteinase with thrombospondin type 1 motif, 8	605175	11q24.3	11	130,280,725	130,410,830			1			EU, AS	GWAS	27,878	Springelkamp et al., 2014 [103]
rs4936099	*ADAMTS8*	a disintegrin-like and metalloproteinase with thrombospondin type 1 motif, 8	605175	11q24.3	11	130,280,725	130,410,830			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	Locus/Gene/ Nearest gene (SNPs)	Gene name	MIM number	Locus	Chr	Hg19 (2009)	Hg38 (2013)	POAG	IOP	VCDR	CCT	ODA or OCA	Population	Study Type	Study size	Reference
rs55796939	*ADAMTS8*	a disintegrin-like and metalloproteinase with thrombospondin type 1 motif, 8	605175	11q24.3	11	130,284,041	130,414,146		1				EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs56009602	*ADAMTS8*	a disintegrin-like and metalloproteinase with thrombospondin type 1 motif, 8	605175	11q24.3	11	130,289,612	130,419,717				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs7481514	*NTM*	neurotrimin	607938	11q25	11	131,291,963	131,422,069	1			1		D:US-Cau AS:US-Cau	GWAS AS	D: 1117 AS: 6469	Ulmer et al., 2012 [196]
rs151326733	*DERA*	deoxyribose-phosphate aldolase	can't find	12p12.3	12	16,189,478	16,036,544	1					D: AA D: HI	GWAS	D: 10,961	Hoffmann et al., 2014 [172]
rs4931170	*FAR2*	fatty acyl CoA reductase 2	616156	12p11.22	12	29,388,772	29,235,839		1				US-Cau	GWAS	D: 1660	Chen et al., 2015 [143]
rs7977237	*LOC105369739 / ADAMTS20*	uncharacterized LOC105369739 / a disintegrin-like and metalloproteinase with thrombospondin type 1 motif 20	-/611681	12q12	12	43,700,341	43,306,538		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs11168187	*RPAP3*	RNA polymerase II-associated protein 3	611477	12q13.11	12	48,044,011	47,650,228			1			EU, AS	GWAS	27,878	Springelkamp et al., 2014 [103]
CNV	*TBK1 (*GLC1P*)*	tank-binding kinase 1	177700	12q14	12	64,845,840 - 64,895,899	64,451,880 - 64,502,108	1 N					AA	LS	1 family	Locus: Fingert et al., 2011 [61] Gene: Fingert et al., 2014 [75]
CNV	*TBK*	tank-binding kinase 1	177700	12q14	12	64,845,840 - 64,895,899	64,451,880 - 64,502,108	N					AU, NZ	CNV	NTG: 334 HTG: 1045 C: 254	Awadalla et al., 2015 [62]
rs343093	*HMGA2*	high mobility group AT-hook 2	600698	12q14.3	12	66,255,005	65,861,225	1					Japan	GWAS	POAG:7378 C:36385	Shiga et al., 2018 [337]
rs7961953	*TMTC2*	transmembrane and tetratricopeptide repeat containing 2	615856	12q21.31	12	83,091,836	82,698,057	1					D: JA R: JA	GWAS	D: 1519 R: 700	Nakano et al., 2009 [76]
rs7961953	*TMTC2*	transmembrane and tetratricopeptide repeat containing 2	615856	12q21.31	12	83,091,836	82,698,057	0					AF	AS	437	Cao et al., 2012 [96]
rs7961953	*TMTC2*	transmembrane and tetratricopeptide repeat containing 2	615856	12q21.31	12	83,091,836	82,698,057	0					CH	AS	1039	Chen et al., 2012b [326]
rs10862688	*TMTC2*	transmembrane and tetratricopeptide repeat containing 2	615856	12q21.31	12	83,922,912	83,529,133			1			EU, AS	GWAS	27,878	Springelkamp et al., 2014 [103]
rs324794	*TMTC2*	transmembrane and tetratricopeptide repeat containing 2	615856	12q21.31	12	83,946,450	83,552,671	1					D: US, L, AS, AA R: multi ethnic	GWAS	D: 63,412 R: 176,890	Choquet et al., 2018 [111]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	**Locus/Gene/ Nearest gene (SNPs)**	**Gene name**	**MIM number**	**Locus**	**Chr**	**Hg19 (2009)**	**Hg38 (2013)**	**POAG**	**IOP**	**VCDR**	**CCT**	**ODA or OCA**	**Population**	**Study Type**	**Study size**	**Reference**
rs442376	*TMTC2*	transmembrane and tetratricopeptide repeat containing 2	615856	12q21.31	12	83,978,117	83,584,338			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs482507	*TMTC2*	transmembrane and tetratricopeptide Repeat containing 2	615856	12q21.31	12	83,979,286	83,585,507	1		1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs324780	*TMTC2*	transmembrane and tetratricopeptide repeat containing 2	615856	12q21.31	12	84,003,866	83,610,087	1		1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs1511589	*TMTC2*	transmembrane and tetratricopeptide repeat containing 2	615856	12q21.31	12	84,061,431	83,667,652					1	EU AS	GWAS	17248 6841	Springelkamp et al., 2015 [134]
rs7972528	*TMTC2*	transmembrane and tetratricopeptide repeat containing 2	615856	12q21.31	12	84,131,036	83,737,257			1			EU AS	GWAS	17248 6841	Springelkamp et al., 2015 [134]
rs7308752	*DCN*	decorin	125255	12q21.33	12	91,527,181	91,133,404				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs11111869	*NT5DC3*	5'-nucleotidase domain containing 3	-	12q23.3	12	104,008,707	104,008,707				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs116878472	*NT5DC3*	5'-nucleotidase domain containing 3	611076	12q23.3	12	104,210,992	103,817,214				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs11553764	*GLT8D2*	glycosyltransferase 8 domain containing 2	-	12q23.3	12	104,415,244	104,021,466				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs1564892	*GLT8D2*	glycosyltransferase 8 domain containing 2	-	12q23.3	12	104,445,742	104,051,964				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
rs74481774	*TMEM119*	transmembrane protein 119	-	12q3.3	12	108,987,230	108,593,454		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs7137828	*ATXN2*	ataxin 2	601517	12q24.12	12	111,932,800	111,494,996	1					Meta-analysis US, AU, EU, SC	GWAS	37333	Bailey et al., 2016 [12]
rs7137828	*ATXN2*	ataxin 2	601517	12q24.12	12	111,932,800	111,494,996	1					US, L, AS, AA	AS	63,412	Choquet et al., 2018 [111]
rs7311936	*FAM101A*	family with sequence similarity 101, member A	615927	12q24.31	12	124,631,597	124,147,051			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs11613189	*FAM101A*	family with sequence similarity 101, member A	615927	12q24.31	12	124,642,803	124,158,257			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs10846617	*FAM101A*	family with sequence similarity 101, member A	615927	12q24.31	12	124,662,131	124,177,585			1			EU AS	GWAS	17248 6841	Springelkamp et al., 2015 [134]
rs1034200	*BASP1P1*	brain abundant, membrane attached signal protein 1 pseudogene 1	-	13q12.12	13	23,228,691	22,654,552				1		D: CR, SCL R: CR, SCL	GWAS	D: 1445 R: 824	Vitart et al., 2010 [201]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	Locus/Gene/ Nearest gene (SNPs)	Gene name	MIM number	Locus	Chr	Hg19 (2009)	Hg38 (2013)	POAG	IOP	VCDR	CCT	ODA or OCA	Population	Study Type	Study size	Reference
rs1034200	*BASP1P1*	brain abundant, membrane attached signal protein 1 pseudogene 1	-	13q12.12	13	23,228,691	22,654,552				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
rs9552680	*BASP1P1*	brain abundant, membrane attached signal protein 1 pseudogene 1	-	13q12.12	13	23,235,285	22,661,146		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs10161679	*BASP1P1*	brain abundant, membrane attached signal protein 1 pseudogene 1	-	13q12.12	13	23,243,645	22,669,506				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
indel	*DCLK1*	doublecortin-like kinase 1	604724	13q13	13	36,629,905	36,055,768			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs7323428	*DCLK1*	doublecortin-like kinase 1	604724	13q13	13	36,643,601	36,069,464			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs1926320	*DCLK1*	doublecortin-like kinase 1	604724	13q13	13	36,652,617	36,078,480			1			D:NL R:NL, UK	GWAS	D: 7360 R: 4455	Ramdas et al., 2010 [99]
rs9546434	*DCLK1*	doublecortin-like kinase 1	604724	13q13	13	36,694,391	36,120,254			1			EU AS	GWAS	17248 6841	Springelkamp et al., 2015 [134]
rs2755237	*FOXO1*	forkhead box O1	136533	13q14.11	13	41,109,429	40,535,292				1		AU, UK	GWAS	D: 5058	Lu et al., 2010 [199]
rs2721051	*FOXO1*	forkhead box O1	136533	13q14.11	13	41,110,884	40,536,747				1		AU, UK	GWAS	D: 5058	Lu et al., 2010 [199]
rs2721051	*FOXO1*	forkhead box O1	136533	13q14.11	13	41,110,884	40,536,747				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
indel	*FOXO1*	forkhead box O1	136533	13q14.11	13	41,112,152	40,538,015				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs11616662	*FOXO1*	forkhead box O1	136533	13q14.11	13	41,119,466	40,545,329		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs9530458	*LMO7*	LIM domain 7	604362	13q22.2	13	76,249,275	75,675,139	1					EU	GWAS	OAG:3071 C:6750	Gharahkhani et al., 2018 [131]
	–			13q31.1	13	79,074,135 - 87,752,255	78,500,000 - 87,100,000		1				BDES	LS	multiple families	Duggal et al., 2005 [188]
	–			14q11.1–14q11.2	14	21,270,988 - 22,787,784	17,200,001 - 24,100,000	1					US	LS	sibpairs	Wiggs et al., 2000 [190]
rs718433	*TRAJ17*	T-cell receptor alpha chain joining gene cluster	615443	14q11.2	14	22,235,890	21,767,650		1				D: EU	GWAS	D: 6236	Ozel et al., 2014 [110]
rs3794453	*RBM23*	RNA binding motif protein 23	–	14q11.2	14	23,386,875	22,917,666			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
indel	*RBM23*	RNA binding motif protein 23	–	14q11.2	14	23,388,793	22,919,584			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
	–			14q21.1–14q21.3	14	37,800,001 - 50,900,000	37,400,001 - 50,400,000	1					US	LS	sibpairs	Wiggs et al., 2000 [190]
rs10130556	*DDHD1/BMP4*	DDHD domain-containing protein 1/ Bone morphogenetic protein 4	614603 / 112262	14q22.1	14	53,970,675	53,503,957			1			EU AS	GWAS	17248 6841	Springelkamp et al., 2015 [134]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	Locus/Gene/ Nearest gene (SNPs)	Gene name	MIM number	Locus	Chr	Hg19 (2009)	Hg38 (2013)	POAG	IOP	VCDR	CCT	ODA or OCA	Population	Study Type	Study size	Reference
rs2251069	*DDHD1/BMP4*	DDHD domain-containing protein 1/ Bone morphogenetic protein 4	614603 / 112262	14q22.1	14	53,988,050	53,521,332			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs4901977	*SIX6*	sine oculis homeobox, drosophila, homolog of, 6	601205 / 606326	14q23.1	14	60,789,176	60,322,458			1			EU, AS	GWAS	27,878	Springelkamp et al., 2014 [103]
rs6573307	*SIX6*	sine oculis homeobox, drosophila, homolog of, 6	606326	14q23.1	14	60,798,009	60,331,291	1					Japan	GWAS	POAG:7378 C:36385	Shiga et al., 2018 [337]
rs4436712	*SIX6*	sine oculis homeobox, drosophila, homolog of, 6	606326	14q23.1	14	60,808,002	60,341,284	1		1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs8015152	*SIX6*	sine oculis homeobox, drosophila, homolog of, 6	606326	14q23.1	14	60,811,999	60,345,281	1		1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs146737847	*SIX6*	sine oculis homeobox, drosophila, homolog of, 6	606326	14q23.1	14	60,976,501	60,509,783	1		1			D: NL, UK R: NL, UK	GWAS	D: 1500 R: 11473	Iglesias et al., 2014 [107]
rs33912345	*SIX6*	sine oculis homeobox, drosophila, homolog of, 6	606326	14q23.1	14	60,976,537	60,509,819	1					US-Cau	AS	518	Carnes et al., 2014 [108]
rs33912345	*SIX6*	sine oculis homeobox, drosophila, homolog of, 6	606326	14q23.1	14	60,976,537	60,509,819	1		1			D: NL, UK R: NL, UK	GWAS	D: 1500 R: 11473	Iglesias et al., 2014 [107]
rs33912345	*SIX6*	sine oculis homeobox, drosophila, homolog of, 6	606326	14q23.1	14	60,976,537	60,509,819	1 N					CH	AS	1132	Sang et al., 2016 [377]
rs33912345	*SIX6*	sine oculis homeobox, drosophila, homolog of, 6	606326	14q23.1	14	60,976,537	60,509,819	1					Meta-analysis US, AU, EU, SC	GWAS	37333	Bailey et al., 2016 [12]
rs10483727	*SIX1/SIX6*	sine oculis homeobox, drosophila, homolog of, 1/6	601205 / 606326	14q23.1	14	61,072,875	60,606,157			1			D: NL R: NL, UK	GWAS	D: 7360 R: 4455	Ramdas et al., 2010 [99]
rs10483727	*SIX1/SIX6*	sine oculis homeobox, drosophila, homolog of, 1/6	601205 / 606326	14q23.1	14	61,072,875	60,606,157			1			D: NL	GWAS	D : 23000	Axenovich et al., 2011 [104]
rs10483727	*SIX1/SIX6*	sine oculis homeobox, drosophila, homolog of, 1/6	601205 / 606326	14q23.1	14	61,072,875	60,606,157	1		1			US-Cau	AS	875	Fan et al., 2011 [97]
rs10483727	*SIX1/SIX6*	sine oculis homeobox, drosophila, homolog of, 1/6	601205 / 606326	14q23.1	14	61,072,875	60,606,157	1					EU	AS	45998	Ramdas et al., 2011 [69]
rs10483727	*SIX1/SIX6*	sine oculis homeobox, drosophila, homolog of, 1/6	601205 / 606326	14q23.1	14	61,072,875	60,606,157	0					AF	AS	437	Cao et al., 2012 [96]
rs10483727	*SIX1/SIX6*	sine oculis homeobox, drosophila, homolog of, 1/6	601205 / 606326	14q23.1	14	61,072,875	60,606,157	1		1			AU, NZ	AS	1759	Dimasi et al., 2012 [29]
rs10483727	*SIX1/SIX6*	sine oculis homeobox, drosophila, homolog of, 1/6	601205 / 606326	14q23.1	14	61,072,875	60,606,157	0		0			D: UK R: UK	GWAS	D: 437 R: 344	Gibson et al., 2012 [93]
rs10483727	*SIX1/SIX6*	sine oculis homeobox, drosophila, homolog of, 1/6	601205 / 606326	14q23.1	14	61,072,875	60,606,157	N		0			JA	AS	616	Mabuchi et al., 2012 [98]
rs10483727	*SIX1/SIX6*	sine oculis homeobox, drosophila, homolog of, 1/6	601205 / 606326	14q23.1	14	61,072,875	60,606,157	1					D: JA R: JA	GWAS	D: 7993 R: 9014	Osman et al., 2012 [90]
rs10483727	*SIX1/SIX6*	sine oculis homeobox, drosophila, homolog of, 1/6	601205 / 606326	14q23.1	14	61,072,875	60,606,157	1		1			D: US-Cau	GWAS	D: 6633	Wiggs et al., 2012 [94]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	Locus/Gene/ Nearest gene (SNPs)	Gene name	MIM number	Locus	Chr	Hg19 (2009)	Hg38 (2013)	POAG	IOP	VCDR	CCT	ODA or OCA	Population	Study Type	Study size	Reference
rs10483727	*SIX1/SIX6*	sine oculis homeobox, drosophila, homolog of, 1/6	601205 / 606326	14q23.1	14	61,072,875	60,606,157	1					D: EU R: EU	GWAS	D: 3077 R: 13044	Gharahkhani et al., 2014 [89]
rs10483727	*SIX1/SIX6*	sine oculis homeobox, drosophila, homolog of, 1/6	601205 / 606326	14q23.1	14	61,072,875	60,606,157	1					BMES	AS	OAG: 67 C: 1919	Burdon et al 2015 [95]
rs10483727	*SIX1/SIX6*	sine oculis homeobox, drosophila, homolog of, 1/6	601205 / 606326	14q23.1	14	61,072,875	60,606,157			1			IN	AS	468	Philomenadin et al., 2015 [105]
rs10483727	*SIX1/SIX6*	sine oculis homeobox, drosophila, homolog of, 1/6	601205 / 606326	14q23.1	14	61,072,875	60,606,157			1			EU AS	GWAS	17248 6841	Springelkamp et al., 2015 [134]
rs10483727	*SIX1/SIX6*	sine oculis homeobox, drosophila, homolog of, 1/6	601205 / 606326	14q23.1	14	61,072,875	60,606,157	1 N					CH	AS	1132	Sang et al., 2016 [377]
rs10483727	*SIX1/SIX6*	sine oculis homeobox, drosophila, homolog of, 1/6	601205 / 606326	14q23.1	14	61,072,875	60,606,157	1					US, L, AS, AA	AS	63,412	Choquet et al., 2018 [111]
rs34935520	*SIX1/SIX6*	sine oculis homeobox, drosophila, homolog of, 1/6	601205 / 606326	14q23.1	14	61,091,401	60,624,683	1		1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs34935520	*SIX1/SIX6*	sine oculis homeobox, drosophila, homolog of, 1/6	601205 / 606326	14q23.1	14	61,091,401	60,624,683	1					D: multi ethnic R: US, L, AS, AA	GWAS	D: 176,890 R: 63,412	Choquet et al., 2018 [111]
rs35155027	*SIX1/SIX6*	sine oculis homeobox, drosophila, homolog of, 1/6	601205 / 606326	14q23.1	14	61,095,174	60,628,456	1					D: US, L, AS, AA R: multi ethnic	GWAS	D: 63,412 R: 176,890	Choquet et al., 2018 [111]
rs56223983	*STON2*	stonin 2	608467	14q31.1	14	81,814,754	81,348,410				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs754203	*CYP46A1*	cytochrome P450, family 46, subfamily A, polypeptide 1	604087	14q32.1	14	100,157,967	99,691,630	1					FR	AS	268	Fourgeux et al., 2009 [378]
rs754203	*CYP46A1*	cytochrome P450, family 46, subfamily A, polypeptide 1	604087	14q32.1	14	100,157,967	99,691,630	1					IN	AS	122	Chandra et al., 2016 [334]
	GLC1I		609745	15q11–15q13	15	19,000,001 - 33,600,000	19,000,001 - 33,400,000	1					US, AA, HI	LS	multiple families	Allingham et al., 2005 [177]
rs785422	*TJP1*	tight junction protein 1	601009	15q13.1	15	30,173,885	29,881,682				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
rs785422	*TJP1*	tight junction protein 1	601009	15q13.1	15	30,173,885	29,881,682				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs28480457	*MEIS2*	MEIS1, mouse, homolog of, 2	601740	15q14	15	37,171,641	36,879,440	1					Japan	GWAS	POAG:7378 C:36385	Shiga et al., 2018 [337]
rs587847	*MEIS2*	MEIS1, mouse, homolog of, 2	601740	15q14	15	37,660,049	37,367,848		1				D: EU	GWAS	D: 6236	Ozel et al., 2014 [110]
rs17352842	*FBN1*	fibrillin 1	134797	15q21.1	15	48,694,211	48,402,014				1		AU	AS	956	Dimasi et al., 2010 [374]
rs8030753	*FBN1*	fibrillin 1	134797	15q21.1	15	48,801,935	48,509,738				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs2593221	*TCF12*	transcription factor 12	600480	15q21.3	15	57,501,414		1					D: US, L, AS, AA R: multi ethnic	GWAS	D: 63,412 R: 176,890	Choquet et al., 2018 [111]
	GLC1N		611274	15q22–15q24	15	59,100,001 - 78,300,000	58,800,001 - 78,000,000	J					CH	LS	1 large family	Wang et al., 2006b [49]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	Locus/Gene/ Nearest gene (SNPs)	Gene name	MIM number	Locus	Chr	Hg19 (2009)	Hg38 (2013)	POAG	IOP	VCDR	CCT	ODA or OCA	Population	Study Type	Study size	Reference
rs12912010	*SMAD3*	SMAD family member 3	603109	15q22.33	15	67,467,143	67,174,805				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs12912045	*SMAD3*	SMAD family member 3	603109	15q22.33	15	67,467,297	67,174,959		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs12913547	*SMAD3*	SMAD family member 3	603109	15q22.33	15	67,467,507	67,175,169				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
rs1048661	*LOXL1*	lysyl oxidase-like 1	153456	15q24.1	15	74,219,546	73,927,205	1					Mixed	AS	5293	Wu et al., 2015 [379]
rs1048661	*LOXL1*	lysyl oxidase-like 1	153456	15q24.1	15	74,219,546	73,927,205	1					Japan	GWAS	POAG:7378 C:36385	Shiga et al., 2018 [337]
rs3825942	*LOXL1*	lysyl oxidase-like 1	153456	15q24.1	15	74,219,582	73,927,241	1					Mixed	AS	5293	Wu et al., 2015 [379]
rs2165241	*LOXL1*	lysyl oxidase-like 1	153456	15q24.1	15	74,222,202	73,929,861	1					Mixed	AS	5293	Wu et al., 2015 [379]
rs2165241	*LOXL1*	lysyl oxidase-like 1	153456	15q24.1	15	74,222,202	73,929,861	1					Spain	AS	658	Zanon-Moreno et al., 2015 [380]
rs6496932	*AKAP13*	A-kinase anchor protein 13	604686	15q25.3	15	85,825,567	85,282,336				1		D: CR, SCL R: CR, SCL	GWAS	D: 1445 R: 824	Vitart et al., 2010 [201]
rs6496932	*AKAP13*	A-kinase anchor protein 13	604686	15q25.3	15	85,825,567	85,282,336				1		D: AS R: AS	GWAS	D: 7711 R: 2681	Cornes et al., 2012 [316]
rs6496932	*AKAP13*	A-kinase anchor protein 13	604686	15q25.3	15	85,825,567	85,282,336				0		D:UK R:UK	GWAS	D: 437 R: 344	Gibson et al., 2012 [93]
rs6496932	*AKAP13*	A-kinase anchor protein 13	604686	15q25.3	15	85,825,567	85,282,336	0			0		D:US-Cau AS:US-Cau	GWAS AS	D: 1117 AS: 6469	Ulmer et al., 2012 [196]
rs6496932	*AKAP13*	A-kinase anchor protein 13	604686	15q25.3	15	85,825,567	85,282,336				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
rs4843040	*AKAP13*	A-kinase anchor protein 13	604686	15q25.3	15	85,838,636	85,295,405				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs1828481	*AKAP13*	A-kinase anchor protein 13	604686	15q25.3	15	85,840,912	85,297,681				1		D: AS R: AS	GWAS	D: 7711 R: 2681	Cornes et al., 2012 [316]
rs1828481	*AKAP13*	A-kinase anchor protein 13	604686	15q25.3	15	85,840,912	85,297,681	0			0		D:US-Cau AS:US-Cau	GWAS AS	D: 1117 AS: 6469	Ulmer et al., 2012 [196]
rs8034595	*NR2F2*	nuclear receptor subfamily 2, group F, member 2	107773	15q26.2	15	96,719,229	96,176,000					1	EU AS	GWAS	17248 6841	Springelkamp et al., 2015 [134]
rs6598351	*FAM169B*	family with sequence similarity 169, member B	-	15q26.3	15	98,808,111	98,264,882			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs72755233	*ADAMTS17*	A disintegrin-like and metalloproteinase with thrombospondin type 1 motif 17	607511	15q26.3	15	100,692,953	100,152,748		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs11247230	*ASB7*	ankyrin repeat- and SOCS box-containing protein 7	615052	15q26.3	15	101,197,005	100,656,800			1			EU AS	GWAS	17248 6841	Springelkamp et al., 2015 [134]
rs60779155	*ASB7*	ankyrin repeat- and SOCS box-containing protein 7	615052	15q26.3	15	101,199,737	100,659,532			1		1	EU AS	GWAS	17248 6841	Springelkamp et al., 2015 [134]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	**Locus/Gene/ Nearest gene (SNPs)**	**Gene name**	**MIM number**	**Locus**	**Chr**	**Hg19 (2009)**	**Hg38 (2013)**	**POAG**	**IOP**	**VCDR**	**CCT**	**ODA or OCA**	**Population**	**Study Type**	**Study size**	**Reference**
rs34222435	*ASB7*	ankyrin repeat- and SOCS box-containing protein 7	615052	15q26.3	15	101,200,873	100,660,668			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs4299136	*ASB7*	ankyrin repeat- and SOCS box-containing protein 7	615052	15q26.3	15	101,201,604	100,661,399			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs2034809	*LRRK1*	leucine-rich repeat kinase 1	610986	15q26.3	15	101,555,399	101,015,194				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
rs2034809	*LRRK1*	leucine-rich repeat kinase 1	610986	15q26.3	15	101,555,399	101,015,194				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs930847	*LRRK1*	leucine-rich repeat kinase 1	610986	15q26.3	15	101,558,562	101,018,357				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
rs930847	*LRRK1*	leucine-rich repeat kinase 1	610986	15q26.3	15	101,558,562	101,018,357				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs4965359	*CHSY1*	chondroitin sulfate synthase 1	608183	15q26.3	15	101,585,336	101,045,131				1		D:AS R:AS	GWAS	D: 7711 R: 2681	Cornes et al., 2012 [316]
rs752092	*CHSY1*	chondroitin sulfate synthase 1	608183	15q26.3	15	101,781,934	101,241,729				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
rs752092	*CHSY1*	chondroitin sulfate synthase 1	608183	15q26.3	15	101,781,934	101,241,729				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs192917960	*RBFOX1*	RNA-binding protein FOX1, C elegans, homolog of, 1	605104	16p13.3	16	5,919,655	5,869,654	1					D: AA D: HI	GWAS	D: 10,961	Hoffmann et al., 2014 [172]
rs192917960	*RBFOX1*	RNA-binding protein FOX1, C elegans, homolog of, 1	605104	16p13.3	16	5,919,655	5,869,654	0					US, L, AS, AA	AS	63,412	Choquet et al., 2018 [111]
rs1906060	*RBFOX1*	RNA-binding protein FOX1, C elegans, homolog of, 1	605104	16p13.3	16	6,108,430	6,058,429		1				D: EU	GWAS	D: 6236	Ozel et al., 2014 [110]
rs3785176	*PMM2*	phosphomannomutase 2	601785	16p13.2	16	8,896,931	8,803,074		1				D: CH R: CH	GWAS	D: 1971 R: 3337	Chen et al., 2014 [366]
rs3785176	*PMM2*	phosphomannomutase 2	601785	16p13.2	16	8,896,931	8,803,074	0					US, L, AS, AA	AS	63,412	Choquet et al., 2018 [111]
rs11646917	*SALL1*	sal-like 1	602218	16q12.1	16	51,428,908	51,394,997			1			EU AS	GWAS	17248 6841	Springelkamp et al., 2015 [134]
rs11646917	*SALL1*	sal-like 1	602218	16q12.1	16	51,428,908	51,394,997			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs1362756	*SALL1*	sal-like 1	602218	16q12.1	16	51,458,290	51,424,379					1	D:NL R:NL, UK	GWAS	D: 7360 R: 4455	Ramdas et al. 2010 [99]
rs1362756	*SALL1*	sal-like 1	602218	16q12.1	16	51,458,290	51,424,379					1	EU AS	GWAS	17248 6841	Springelkamp et al., 2015 [134]
indel	*SALL1*	sal-like 1	602218	16q12.1	16	51,461,915	51,428,004	1		1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs4784295	*SALL1*	sal-like 1	602218	16q12.1	16	51,471,432	51,437,521			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs1345467	*SALL1*	sal-like 1	602218	16q12.1	16	51,482,321	51,448,410			1			EU, AS	GWAS	27,878	Springelkamp et al., 2014 [103]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	Locus/Gene/ Nearest gene (SNPs)	Gene name	MIM number	Locus	Chr	Hg19 (2009)	Hg38 (2013)	POAG	IOP	VCDR	CCT	ODA or OCA	Population	Study Type	Study size	Reference
rs1345467	*SALL1*	sal-like 1	602218	16q12.1	16	51,482,321	51,448,410	1		1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
3'UTR	*CDH-1*	cadherin 1	192090	16q22.1	16	68,867,398 - 68,869,440	68,833,495 - 68,835,537	1					CH	AS	163	Lin et al., 2006 [381]
rs75828804	*ADAMTS18 / NUDT7*	A disintegrin-like and metalloproteinase with thrombospondin type 1 motif 18 / nudix hydrolase 7	607512/609231	16q23.1	16	77,591,935	77,558,038		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs12447690	*ZNF469*	zinc finger protein 469	612078	16q24.2	16	88,298,124	88,264,518				1		D: AU, UK	GWAS	D: 5058	Lu et al., 2010 [199]
rs12447690	*ZNF469*	zinc finger protein 469	612078	16q24.2	16	88,298,124	88,264,518				1		D: CR, SCL R: CR, SCL	GWAS	D: 1445 R: 824	Vitart et al., 2010 [201]
rs12447690	*ZNF469*	zinc finger protein 469	612078	16q24.2	16	88,298,124	88,264,518				1		D1: SM D2: SG-CH	GWAS	D1: 3280 D2: 3400	Vithana et al., 2011 [315]
rs12447690	*ZNF469*	zinc finger protein 469	612078	16q24.2	16	88,298,124	88,264,518				1		D: AS R: AS	GWAS	D: 7711 R: 2681	Cornes et al., 2012 [316]
rs12447690	*ZNF469*	zinc finger protein 469	612078	16q24.2	16	88,298,124	88,264,518				0		AU, NZ	AS	1759	Dimasi et al., 2012 [29]
rs12447690	*ZNF469*	zinc finger protein 469	612078	16q24.2	16	88,298,124	88,264,518				1		D: GE R: NL	GWAS	D: 3931 R: 1418	Hoehn et al., 2012 [198]
rs12447690	*ZNF469*	zinc finger protein 469	612078	16q24.2	16	88,298,124	88,264,518				1		D:US-Cau AS:US-Cau	GWAS AS	D: 1117 AS: 6469	Ulmer et al., 2012 [196]
rs6540223	*ZNF469*	zinc finger protein 469	612078	16q24.2	16	88,321,436	88,287,830				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
rs35193497	*ZNF469*	zinc finger protein 469	612078	16q24.2	16	88,324,821	88,291,215				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs28687756	*ZNF469*	zinc finger protein 469	612078	16q24.2	16	88,328,928	88,295,322				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs12926024	*ZNF469*	zinc finger protein 469	612078	16q24.2	16	88,331,309	88,297,703		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs9938149	*ZNF469*	zinc finger protein 469	612078	16q24.2	16	88,331,640	88,298,034				1		D: AU, UK	GWAS	D: 5058	Lu et al., 2010 [199]
rs9938149	*ZNF469*	zinc finger protein 469	612078	16q24.2	16	88,331,640	88,298,034				1		D1: SM D2: SG-CH	GWAS	D1: 3280 D2: 3400	Vithana et al., 2011 [315]
rs9938149	*ZNF469*	zinc finger protein 469	612078	16q24.2	16	88,331,640	88,298,034				0		AU, NZ	AS	1759	Dimasi et al., 2012 [29]
rs9938149	*ZNF469*	zinc finger protein 469	612078	16q24.2	16	88,331,640	88,298,034				1		D: GE R: NL	GWAS	D: 3931 R: 1418	Hoehn et al., 2012 [198]
rs9938149	*ZNF469*	zinc finger protein 469	612078	16q24.2	16	88,331,640	88,298,034				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
	–			17p13.3 - 17p13.1	17	1 - 10,700,000	1 - 10,800,000	1					US	LS	sibpairs	Wiggs et al., 2000 [190]
rs4790881	*SMG6*	SMG6, nonsense mediated mRNA decay factor	610963	17p13.3	17	2,068,932	2,165,638		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	**Locus/Gene/ Nearest gene (SNPs)**	**Gene name**	**MIM number**	**Locus**	**Chr**	**Hg19 (2009)**	**Hg38 (2013)**	**POAG**	**IOP**	**VCDR**	**CCT**	**ODA or OCA**	**Population**	**Study Type**	**Study size**	**Reference**
rs34629349	*INCA1*	inhibitor of CDK, cyclin A1 interacting protein 1	617374	17p13.2	17	4,829,560	4,926,265		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs1042522	*TP53*	tumour protein p53	191170	17p13.1	17	7,579,472	7,676,154	1					CH	AS	117	Lin et al., 2002 [261]
rs1042522	*TP53*	tumour protein p53	191170	17p13.1	17	7,579,472	7,676,154	1					IR	AS	130	Neamatzadeh et al., 2015 [262]
rs1042522	*TP53*	tumour protein p53	191170	17p13.1	17	7,579,472	7,676,154	N					CH	AS	606	Fan et al., 2010 [264]
rs1042522	*TP53*	tumour protein p53	191170	17p13.1	17	7,579,472	7,676,154	1					Cau US	AS	358	Daugherty et al., 2009 [263]
rs9897123	*GAS7*	growth arrest-specific 7	603127	17p13.1	17	10,020,501	10,117,184	1					US, L, AS, AA	AS	63,412	Choquet et al., 2018 [111]
rs9897123	*GAS7*	growth arrest-specific 7	603127	17p13.1	17	10,020,501	10,117,184	1					Meta-analysis US, AU, EU, SC	GWAS	37333	Bailey et al., 2016 [12]
rs12150284	*GAS7*	growth arrest-specific 7	603127	17p13.1	17	10,031,090	10,127,773		1				D: EU	GWAS	D: 6236	Ozel et al., 2014 [110]
rs9913911	*GAS7*	growth arrest-specific 7	603127	17p13.1	17	10,031,183	10,127,866		1				D: AS D: EU	GWAS	D: 7738 D:27558	Hysi et al., 2014 [86]
rs9913911	*GAS7*	growth arrest-specific 7	603127	17p13.1	17	10,031,183	10,127,866	1	1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs9913911	*GAS7*	growth arrest-specific 7	603127	17p13.1	17	10,031,183	10,127,866	1	1	1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs9913911	*GAS7*	growth arrest-specific 7	603127	17p13.1	17	10,031,183	10,127,866	1	1				D: US, L, AS, AA R: multi ethnic	GWAS	D: 63,412 R: 176,890	Choquet et al., 2018 [111]
rs9913911	*GAS7*	growth arrest-specific 7	603127	17p13.1	17	10,031,183	10,127,866	1					D: multi ethnic R: US, L, AS, AA	GWAS	D: 176,890 R: 63,412	Choquet et al., 2018 [111]
rs11656696	*GAS7*	growth arrest-specific 7	603127	17p13.1	17	10,033,679	10,130,362	1	1				D: NL R: UK, AU, NZ, Canada AS: NL, GE	GWAS AS	D: 11972 R: 7482 AS: 1432	van Koolwijk et al., 2012 [66]
rs2323457	*HS3ST3B1*	heparan sulfate-glucosamine 3-sulfotransferase 3B1	604058	17p12	17	14,554,190	14,650,873				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
rs12940030	*HS3ST3B1*	heparan sulfate-glucosamine 3-sulfotransferase 3B1	604058	17p12	17	14,561,016	14,561,016				1		EU, AS	MA	>20,000	Lu et al., 2013 [200]
rs4792535	*HS3ST3B1*	heparan sulfate-glucosamine 3-sulfotransferase 3B1	604058	17p12	17	14,565,130	14,661,813				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs11870935	*KPNB1*	karyopherin beta-1	602738	17q21.32	17	45,732,605	47,655,239			1			EU AS	GWAS	17248 6841	Springelkamp et al., 2015 [134]
rs11867840	*BCAS3*	breast carcinoma amplified sequence 3	607470	17q23	17	59,273,265	61,195,904	1		1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs11651885	*BCAS3*	breast carcinoma amplified sequence 3	607470	17q23	17	59,286,263	61,208,902			1			EU AS	GWAS	17248 6841	Springelkamp et al., 2015 [134]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	**Locus/Gene/ Nearest gene (SNPs)**	**Gene name**	**MIM number**	**Locus**	**Chr**	**Hg19 (2009)**	**Hg38 (2013)**	**POAG**	**IOP**	**VCDR**	**CCT**	**ODA or OCA**	**Population**	**Study Type**	**Study size**	**Reference**
rs8068952	*BCAS3*	breast carcinoma amplified sequence 3	607470	17q23	17	59,286,644	61,209,283			1			D:NL R:NL, UK	GWAS	D:7360 R:4455	Ramdas et al., 2010 [99]
	–			17q25.1 – 17q25.3	17	70,900,001 - 81,195,210	72,900,001 - 83,257,441	1					US	LS	sibpairs	Wiggs et al., 2000 [190]
rs52809447	*GGA3*	golgi-associated, gamma-adaptin ear-containing, ARF-binding protein 3	606006	17q25.1	17	73,238,508	75,242,427		1				US-Cau	GWAS	D: 1660	Chen et al., 2015 [143]
rs11659764	*TCF4*	transcription factor 4	602272	18q21.2	18	53,335,512	55,668,281		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
	–			19p13.2	19	6,900,000 - 13,900,000	6,900,000 - 12,600,000		1				BDES	LS		Duggal et al., 2007 [191]
c.431G→A Arg144Gln	*OLFM2*	olfactomedin 2	–	19p13.2	19	9,968,088	9,857,412	1					JA	AS	770	Funayama et al., 2006 [382]
	–			19q12	19	28,600,001 - 32,400,000	28,100,001 - 31,900,000	1					US	LS	sibpairs	Wiggs et al., 2000 [190]
rs25487	*XRCC1*	X-ray cross-complementing group 1	194360	19q13.31	19	44,055,726	43,551,574	1					PA	AS	516	Yousaf et al., 2011 [383]
rs449647	*APOE*	apolipoprotein E	107741	19q13.2	19	45,408,564	44,905,307	1					PO	AS	769	Nowak et al., 2015 [351]
rs429358 rs7412	*APOE*	apolipoprotein E	107741	19q13.2	19	45,411,941 45,412,079	44,908,684 44,908,822	1					AU, NZ	AS	193	Vickers et al., 2002 [384]
rs429358 rs7412	*APOE*	apolipoprotein E	107741	19q13.2	19	45,411,941 45,412,079	44,908,684 44,908,822	0					UK	AS	504	Lake et al., 2004 [385]
rs429358 rs7412	*APOE*	apolipoprotein E	107741	19q13.2	19	45,411,941 45,412,079	44,908,684 44,908,822	0					UK	AS	212	Ressiniotis et al., 2004 [386]
rs429358 rs7412	*APOE*	apolipoprotein E	107741	19q13.2	19	45,411,941 45,412,079	44,908,684 44,908,822	1					JA	AS	681	Fan et al., 2005 [387]
rs429358 rs7412	*APOE*	apolipoprotein E	107741	19q13.2	19	45,411,941 45,412,079	44,908,684 44,908,822	1	1				JA	AS	489	Mabuchi et al., 2005 [388]
rs429358 rs7412	*APOE*	apolipoprotein E	107741	19q13.2	19	45,411,941 45,412,079	44,908,684 44,908,822	1					CH	AS	700	Lam et al., 2006 [389]
rs429358 rs7412	*APOE*	apolipoprotein E	107741	19q13.2	19	45,411,941 45,412,079	44,908,684 44,908,822	0					ES	AS	429	Zetterberg et al., 2007 [390]
rs429358 rs7412	*APOE*	apolipoprotein E	107741	19q13.2	19	45,411,941 45,412,079	44,908,684 44,908,822	1					SA	AS	190	Al-Dabbagh et al., 2009 [391]
rs429358 rs7412	*APOE*	apolipoprotein E	107741	19q13.2	19	45,411,941 45,412,079	44,908,684 44,908,822	0					TU	AS	194	Saglar et al., 2009 [392]
rs13181	*XPD*	xeroderma pigmentosum complementation group D	278730	19q13.32	19	45,854,919	45,351,661	1					PA	AS	516	Yousaf et al., 2011 [383]
	*NTF4 (*GLC1O*)*	neurotrophin-4	162662	19q13.3	19	49,558,618 - 49,568,333	49,061,066 - 49,061,997	1 N					GE, NL	AS	1,379	Pasutto et al., 2009 [74]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	**Locus/Gene/ Nearest gene (SNPs)**	**Gene name**	**MIM number**	**Locus**	**Chr**	**Hg19 (2009)**	**Hg38 (2013)**	**POAG**	**IOP**	**VCDR**	**CCT**	**ODA or OCA**	**Population**	**Study Type**	**Study size**	**Reference**
rs3116139	*NKG7*	natural killer cell group 7 sequence	606008	19q13.41	19	51,879,034	51,375,780	1					D: AA D: HI	GWAS	D: 10,961	Hoffmann et al., 2014 [172]
rs1279683	*SLC23A2*	solute carrier family 23, member 2	603791	20p13	20	4,983,092	5,002,446	1					ME	AS	500	Zanon-Moreno et al., 2013 [393]
	GLC1K		608696	20p12	20	5,100,001 - 17,900,000	5,100,001 - 17,900,000	J					US	LS	several families	Wiggs et al., 2004 [47]
rs6054374	*BMP2*	bone morphogenic protein 2	112261	20p12.3	20	6,578,556	6,597,909			1			EU, AS	GWAS	27,878	Springelkamp et al., 2014 [103]
rs6054375	*BMP2*	bone morphogenic protein 2	112261	20p12.3	20	6,578,629	6,597,982	1		1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs6107845	*BMP2*	bone morphogenic protein 2	112261	20p12.3	20	6,578,741	6,598,094	1		1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs3918508	*LINC01734/ LINC01370*	long intergenic non-protein coding RNA 1734 / long intergenic non-protein coding RNA 1370		20q12	20	37,912,667	39,284,024		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs2285142	*PTPRT*	protein-tyrosine phosphatase, receptor type T	608712	20q12	20	40,978,940	42,350,300		1				NIES	GWAS (Ped.)	330	Matovinovic et al., 2017 [133]
rs2839082	*COL6A1/ COL6A2*	collagen type VI alpha 1 chain / collagen type VI alpha 2	120220 / 120240	21q22.3	21	47,442,334	46,022,420		1				D: US, LA, AS, AA R: EU, AS	GWAS	D: 69,756 R: 37,930	Choquet et al., 2017 [171]
rs8133436	*COL6A2*	collagen type VI alpha 2	120240	21q22.3	21	47,519,535	46,099,621				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
rs58714937	*TXNRD2*	thioredoxin reductase 2	606448	22q11.21	22	19,856,710	19,869,187	1					D: multi ethnic R: US, L, AS, AA	GWAS	D: 176,890 R: 63,412	Choquet et al., 2018 [111]
rs35934224	*TXNRD2*	thioredoxin reductase 2	606448	22q11.21	22	19,872,645	19,885,122	1					US, L, AS, AA	AS	63,412	Choquet et al., 2018 [111]
rs35934224	*TXNRD2*	thioredoxin reductase 2	606448	22q11.21	22	19,872,645	19,885,122	1					Meta-analysis US, AU, EU, SC	GWAS	37333	Bailey et al., 2016 [12]
rs71313931	*ARVCF*	armadillo repeat gene deleted in velocardiofacial syndrome	602269	22q11.21	22	19,960,184	19,972,661				1		Meta-analysis EU, AS	GWAS	25,910	Iglesias et al., 2018 [179]
null	*GSTT1*	glutathione S transferase theta 1	600436	22q11.23	22	24,376,139 - 24,384,284	17,400,000 - 25,500,000	1					TU	AS	265	Unal et al., 2007 [322]
rs1547014	*CHEK2*	checkpoint kinase 2	604373	22q12.1	22	29,100,711	28,704,723			1			D: NL R: NL, UK	GWAS	D: 7360 R: 4455	Ramdas et al., 2010 [99]
rs1547014	*CHEK2*	checkpoint kinase 2	604373	22q12.1	22	29,100,711	28,704,723	0					EU	AS	45998	Ramdas et al., 2011 [69]
rs1547014	*CHEK2*	checkpoint kinase 2	604373	22q12.1	22	29,100,711	28,704,723			0			AU, NZ	AS	1759	Dimasi et al., 2012 [29]
rs1547014	*CHEK2*	checkpoint kinase 2	604373	22q12.1	22	29,100,711	28,704,723			1			JA	AS	616	Mabuchi et al., 2012 [98]
rs1547014	*CHEK2*	checkpoint kinase 2	604373	22q12.1	22	29,100,711	28,704,723			1			D: US-Cau	GWAS	D: 6633	Wiggs et al., 2012 [94]

				Genomic location		Chr. position (bp)			Clinical trait				Study characteristics			
SNP/Indel	Locus/Gene/ Nearest gene (SNPs)	Gene name	MIM number	Locus	Chr	Hg19 (2009)	Hg38 (2013)	POAG	IOP	VCDR	CCT	ODA or OCA	Population	Study Type	Study size	Reference
rs1547014	*CHEK2*	checkpoint kinase 2	604373	22q12.1	22	29,100,711	28,704,723			1			EU, AS	GWAS	27,878	Springelkamp et al., 2014 [103]
rs5762752	*CHEK2*	checkpoint kinase 2	604373	22q12.1	22	29,100,977	28,704,989	1		1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs5752773	*CHEK2*	checkpoint kinase 2	604373	22q12.1	22	29,105,415	28,709,427	1		1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs738722	*CHEK2*	checkpoint kinase 2	604373	22q12.1	22	29,130,012	28,734,024	1		1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs1033667	*CHEK2*	checkpoint kinase 2	604373	22q12.1	22	29,130,300	28,734,312			1			EU AS	GWAS	17248 6841	Springelkamp et al., 2015 [134]
rs2412970	*HORMAD2*	HORMA domain containing 2	-	22q12.2	22	30,486,826	30,090,837					1	EU AS	GWAS	17248 6841	Springelkamp et al., 2015 [134]
rs737723	*SEC14L2*	SEC14-like 2	607558	22q12.2	22	30,802,029	30,406,040	1					ME	AS	500	Zanon-Moreno et al., 2013 [393]
rs209217	*CARD10*	caspase recruitment domain-containing protein 10	607209	22q13.1	22	37,907,069	37,511,062			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs56385951	*CARD10*	caspase recruitment domain-containing protein 10	607209	22q13.1	22	37,909,539	37,513,532			1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs9607469	*CARD10*	caspase recruitment domain-containing protein 10	607209	22q13.1	22	37,919,267	37,523,260	0					AF	AS	437	Cao et al., 2012 [96]
rs9607469	*CARD10*	caspase recruitment domain-containing protein 10	607209	22q13.1	22	37,919,267	37,523,260					1	D:AS R:NL	GWAS	D: 4445 R: 9326	Khor et al., 2011 [317]
rs9607469	*CARD10*	caspase recruitment domain-containing protein 10	607209	22q13.1	22	37,919,267	37,523,260				0		D:UK R:UK	GWAS	D: 437 R: 344	Gibson et al., 2012 [93]
rs9607469	*CARD10*	caspase recruitment domain-containing protein 10	607209	22q13.1	22	37,919,267	37,523,260			1			IN	AS	468	Philomenadin et al., 2015 [105]
rs9607469	*CARD10*	caspase recruitment domain-containing protein 10	607209	22q13.1	22	37,919,267	37,523,260					1	EU AS	GWAS	17248 6841	Springelkamp et al., 2015 [134]
rs1074407	*TRIOBP*	TRIO- and F-actin binding protein	609761	22q13.1	22	38,156,183	37,760,176	1		1		1	EU, AS	MA	IGGC >79,000	Springelkamp et al., 2017 [136]
rs5756813	*TRIOBP*	TRIO- and F-actin binding protein	609761	22q13.1	22	38,175,477	37,779,470			1			EU, AS	GWAS	27,878	Springelkamp et al., 2014 [103]
rs5756813	*TRIOBP*	TRIO- and F-actin binding protein	609761	22q13.1	22	38,175,477	37,779,470			1			EU AS	GWAS	17248 6841	Springelkamp et al., 2015 [134]
rs7291444	*PKDREJ*	polycystin and sea urchin REJ homolog-like	604670	22q13.31	22	46,656,246	46,260,349		1				US-Cau	GWAS	D: 1660	Chen et al., 2015 [143]

Appendix C Scripts for MGA

C.1 Subset vcf file and join to MGA data file

C.1.1 Subset vcf file to same regions as extracted from MGA file

bed file used for this is identical to bed file for extracting mgassoc results except no column headings in the bed file here.

```
module load bcftools
bcftools view -R BED_FILE_for_vcf.txt -O v -o
extracted_vcf_section_SNPs_.vcf full_vcf_file.vcf.gz
```

C.1.2 Join extracted MGA results to extracted and annotated vcf files

```
library(tidyr)
library(dplyr)
```

Read in extracted mgassoc file

trait_mgassoc <- read.csv(file = "extracted_mgassoc_file.csv", header = TRUE, sep = ",")

Need to make sure that the bp column is read as numbers

```
trait_mgassoc$bp <- as.numeric(as.character(trait_mgassoc$bp))
```

to remove the chr and bp columns which aren't needed anymore - although sometimes handy to leave for data sorting.

```
trait_final <- select(trait_mgassoc, -Chr, -bp)
```

to join on to annotated, peaks extracted vcf file (NOTE: vcf file needs prep work before this stage)

```
annotatedVCFfile <- read.table("annotatedVCFfile.txt", header = TRUE, sep =
"\t", as.is = TRUE, quote="")
```

to turn it into a table

```
annotatedVCFfile_table <- tbl_df(annotatedVCFfile)
```

to combine columns to make it look like the mgassoc file

need to join by POS rather than Start as indels become "normalised" once annotated, affecting the starting positions of deletions)

some deletions also in SNP file, so joining by POS is better for both SNPs and indels

```
combined_columns_annotatedVCFfile <- unite(annotatedVCFfile_table, SNP, Chr, POS, sep = "_")
```

To combine the mgassoc file with annotated vcf file (mgassoc on left)

```
joined <- inner_join(trait_mgassoc, combined_columns_annotatedVCFfile, by = "SNP")
```

To write the final, joined table as a new file

```
write.csv(joined,
          file = "output_file_mgassoc_annotated.vcf.csv",
          row.names = FALSE)
```

Appendix D Linkage analysis scripts

D.1 R script for removing rare variants and setting low quality calls to "missing"

Using full vcf file, NOT annotated

```
library(tidyr)
library(dplyr)
```

Read in vcf

```
full_vcf <- read.table(file = "full_vcf_file.vcf",
            sep="\t", header=FALSE, quote = "", as.is=TRUE,na.strings=".",
            stringsAsFactors = FALSE, colClasses = "character")
```

Set "Low_Confidence", "Low_coverage" and "Low_quality" as missing (./.)

```
full_vcf_cleaned <-
as.data.frame(sapply(full_vcf,gsub,pattern=".*Low.*",replacement="\\./\\.")
)
```

Count the number of heterozygous variant calls each variant in the file and add this as a new column in the vcf file

The ^ is to only count the genotypes at the beginning of the each cell (the info column has "unwanted" genotypes in it)

The [/|] is to pick up both unphased and phased genotypes. Prob won't make a difference here, but might for future runs

```
full_vcf_cleaned$het_counts <- apply(full_vcf_cleaned, 1, function(x)
sum(grepl('^1[/|]0|^0[/|]1', x)))
```

Count the number of homozygote variant calls each variant in the file and add this as a new column in the vcf file

```
full_vcf_cleaned$hom_counts <- apply(full_vcf_cleaned, 1, function(x)
sum(grepl('^1[/|1]', x)))
```

Create another column which adds 1 per het count and 2 per homozygous count

```
full_vcf_cleaned$total_counts <- (full_vcf_cleaned$het_counts +
2*full_vcf_cleaned$hom_counts)
```

to check whether everything is looking ok

```
head <- full_vcf_cleaned[1:200,]
write.table(head, file = "head.txt",sep = "\t", row.names = F
col.names = F, quote = F,na = ".")
```

Filter vcf file to get rid of rare or too common SNPs (filter is a dplyr function, so make sure it's loaded)

```
filtered_between <- filter(full_vcf_cleaned, (full_vcf_cleaned$total_counts
>= 8 & full_vcf_cleaned$total_counts <= 490))
```

to check whether everything is looking ok

```
top <- filtered_between[1:200,]
write.table(top, file = "top.txt",sep = "\t", row.names = F,
col.names = T, quote = F, na = ".")
```

After filtering, delete the 3 columns which I added in to do the counts (columns 259 to 261)

```
final_vcf <- filtered_between[, -c(259:261)]
```

If this file is going to be used to blank out PedCheck error SNPs, need to have some form of ID no. in ID column

Just copy bp position from column 2 to column 3 to use as ID

```
final_vcf$V3 <- final_vcf$V2
```

Write table for remerging with the vcf header

```
write.table(final_vcf,file = output_cleaned_filtered_vcf_file.vcf",
            sep = "\t", row.names = F, col.names = F, quote = F,
            na = ".")
```

D.2 PLINK command to remove SNPs with missing genotype calls, deviation from Hardy-Weinberg Equilibrium, presence on sex chromosomes and recode as a vcf file.

--geno 0.05 removes variants with missing call rates exceeding 5%
--hwe 0.001 removes variants with a Hardy-Weinberg exact test p-value less than 0.001
--chr 1-22 only retains autosomal variants
--recode vcf creates an output vcf file (necessary for Mendelian error checking steps)

```
plink --vcf input_vcf_file.vcf --geno 0.05 --hwe 0.001 --chr 1-22 \
--recode vcf --out output_vcf_file
```

D.3 File preparation for PedCheck

D.3.1 Subset filtered vcf into individual chromosomes (loop command)

Filtered vcf input file: filteredSNPs.vcf

```
module load bcftools \
for chr in {1..22}; do bcftools view -t chr${chr} -O v -o
chr${chr}filteredSNPs.vcf filteredSNPs.vcf & done
```

Output files: chr${chr}filteredSNPs vcf files

D.3.2 Create PLINK ped and map files (loop command)

alleles needed numerical coding instead of alphabetic

```
module load plink \
for chr in {1..22}; \
do plink --vcf chr${chr}filteredSNPs.vcf --allele1234 --recode --cm-map \
file_path_to/genetic_map_chr@_combined_b37.txt --out chr${chr}PLINKPed &
done
```

Output files: chr$chrPLINKped.map and chr${chr}PLINKPed.ped

D.3.3 Modification of map file for PedCheck (loop command)

concatenate individual chromosome map files into 1 full map file

```
for chr in {1..22}; do cat chr${chr}PLINKPed.map >> all_chr_PLINKPed.map \
& done
```

use excel to concatenate chromosome and base pair position columns into 1 column (chr_bp).

delete all other columns

subset this full file back to individual chromosomes

```
awk 'BEGIN {FS = "_"}; {print > ("chr"$1"NAMESforPedCheck.txt")}' \
NAMESforPedCheck.txt
```

D.3.4 Modification of ped file for PedCheck (loop command)

R script (loop command) for joining PLINK produced ped files with extra pedigree information

```
library(dplyr)
```

Read in ped file (family info only).This has the new ID numbers and is sorted by generations within families.

Changed .ped to .txt before doing this.

```
familyPED <- read.csv(file = "PEDforPedCheck.ped.txt", header = TRUE, sep =
"\t")
```

Start of loop command

```
for (chr in 1:22) {
```

Read in PLINK.ped output file.Added a .txt to the end of the output file.

```
PLINKped <- read.csv(file = paste("chr", chr, "PLINKPed.ped.txt", sep =
''),
FALSE, sep = " ")
```

Name column 2 in the PLINKped file as "ID" (as can't join the 2 files without named column)

```
colnames(PLINKped)[2] <- "ID"
```

Join original ped file to PLINK output ped file. Inner join only retains the rows in both sets (ie only the people with WES info)

```
JoinedPED <- left_join (familyPED, PLINKped, by = "ID")
```

Delete columns not necessary for PedCheck (ie generation, ID,SEX, V1 to V6)

```
FinalPED <- select(JoinedPED, -6,-7,-V1,-V3,-V4,-V5,-V6)
```

write the output as a text file, so that I can manipulate for PedCheck

```
write.table(FinalPED, file = paste("chr", chr, "PEDforPedCheck.ped.txt",
sep = ''),
     sep = "\t",
     quote = FALSE,
     na = "0",
     row.names=FALSE,
     col.names=FALSE)
}
```

D.3.5 Mendelian error checking using PedCheck (loop command)

Loop which runs each chromosome separately (otherwise PedCheck crashes)

As PedCheck outputs an error file with the same name for each chromosome, need to rename the output file before starting on the next chromosome. Run in sequence rather than in parallel for this reason.

```
module load pedcheck
for chr in {1..22}; do \
pedcheck -p chr${chr}PEDforPedCheck.ped \
-n chr${chr}NAMESforPedCheck.txt -4 -e -a; \
mv pedcheck.err chr${chr}PedCheck.err; done
```

D.3.6 Remove Mendelian error SNPs from vcf file (loop command)

SNPs to exclude files made for each chromosome. Just a list of the bp positions in a text file.

module load plink

```
for chr in {1..22}; \
do plink --vcf chr${chr}filteredSNPs.vcf --exclude
chr${chr}_excludeSNPs.txt \
--recode vcf --out chr${chr}PLINKno_errors & done
```

D.4 File preparation for IBDLD

D.4.1 PLINK command for IBDLD file prep

default letter coding for alleles needed for this

```
module load plink
plink --vcf input.vcf --recode --cm-map \
file_path_to/genetic_map_chr@_combined_b37.txt --out output_file_prefix
```

D.4.2 Map file requirements for IBDLD

PLINK output map file needs an additional unique SNP identification number in the second column. Rather than using rs numbers (which not every variant has), I created a unique chromosome_base-pair-location number using “concatenate” with Excel.

Output files: MAP_file_for_IBDLD.map

D.4.3 Modification of ped file for IBDLD

R script to modify PLINK ped file output for IBDLD input

```
library(dplyr)
```

Read in ped file (family info only).This has the new ID numbers and is sorted by generations within families. Changed .ped to .txt before doing this.

```
familyPED <- read.csv(file = "ped_file.ped.txt", header = TRUE, sep = "\t")
```

Read in PLINK.ped output file Added a .txt to the end of the output file.

The colClasses command is to stop R reading in a column of "T" as "True"

```
PLINKped <- read.csv(file = "PLINKped_output_file.ped.txt", header = FALSE,
            sep = " ", colClasses = "character")
```

Name column 2 in the PLINKped file as "ID" (as can't join the 2 files without named column)

```
colnames(PLINKped)[2] <- "ID"
```

Join original ped file to PLINK output ped file. Inner join only retains the rows in both sets (ie only the people with WES info)

```
JoinedPED <- inner_join (familyPED, PLINKped, by = "ID")
```

Delete columns not necessary for IBDLD (ie ID,SEX, V1 to V6)

```
FinalPED <- select(JoinedPED, -6,-V1,-V3,-V4,-V5,-V6)
```

write the output as a text file, so that IBDLD can work with it

```
write.table(FinalPED, file = "PED_file_for_IBDLD.ped ",
      sep = "\t",
      quote = FALSE,
      row.names=FALSE,
      col.names=FALSE)
```

#Output file: PED_file_for_IBDLD.ped

D.4.4 IBDLD command

-ploci 50 -dist 2 uses 50 previous loci withing a 2cM window to calculate LD

-mincallrate 0.95 sets the minimum threshold for each SNP for inclusion in LD calculations and IBD estimation at 95%

-step 0 executes both LD calculation and IBD estimation steps in order

-ibd 2 outputs IBD sharing at each locus and Δ7 coefficient required for SOLAR

-hiddenstates 9 is the appropriate setting for related individuals in a pedigree and sets the IBD hidden states number to 9

--ibdtxt outputs ibd probabilities as a text file

```
module load IBDLD
ibdld -o prefix_for_IBDLD_output_files -p PED_file_for_IBDLD.ped \
-m MAP_file_for_IBDLD.map \
-ploci 50 -dist 2 -mincallrate 0.95 -step 0  -ibd 2 -hiddenstates 9 --
ibdtxt
```

D.5 Generation of Multipoint identity by descent (MIBD) matrices

D.5.1 To create MIBD files from ibd.txt. output files

awk script written by Juan Peralta

#!/usr/bin/env awk -f

build solar mibd files directly from ibdld txt files

pass:

-v prefix="solar/mibd/mibd."

-v indexout="path/to/pedindex.out"

-v indexcde="path/to/pedindex.cde"

-v map="cm_map.tab" (a map with the subset of markers to be extracted into the mibds)

```
BEGIN {
  seen[-1]=1
  addcrc=0
  if (prefix == "") prefix="mibd/mibd."
  if (indexout == "") indexout="pedindex.out"
  if (indexcde == "") indexcde="pedindex.cde"
  col = 1
  len = 1
  FS=" "
  while(getline < indexcde) {
    if (NR==1) continue;
    if ($2=="ID") { len = $1; break; }
    col = col + $1
  }
  while(getline < indexout){
    id=substr($0,col,len)
    gsub(" ","",id)
    ibdid[id]=$1
  }
  FS="\t"
  while(getline < map){
    mchr[$2] = $1
    mcm[$2] = $4
  }

}
```

map marker names to column numbers, 2 cols per name

```
NR == 1 {
  colspermarker = $1
  chr = $3
  j=5
  for (i=5;i<=NF;i++){
    idx[$i]=j
    j=j+colspermarker
  }
```

```
  for (m in mcm) {
    if (!(m in idx)) delete mcm[m]
  }
```

mark mibd with crc from pedindex file

```
  if (! (cm in seen) && addcrc) {
    printf " %5i  %5i  %s %s\n",ibdid1,ibdid2,"0.1923398118",".28249" >>
prefix""$1"."cm
    seen[cm]=1
  }
}
```

ignore famids in the output

save each col into a different mibd file

```
NR > 1{
  for (m in mcm){
    cm = int(mcm[m])
    i = idx[m]
    j = i + 1;
```

#print m,cm,idx[m],i,j >> prefix""chr"."cm

```
    if ($2 == $4) done[$2] = y
    printf "%5i %5i  %.7f  %0.7f\n",ibdid[$2],ibdid[$4],$i*2,$j >>
prefix""chr"."cm
  }
#  if (ibdid1 == ibdid2) {
#      diag[$1"@"cm"@"ibdid1]=1
#  }
#  if (! ($1 in chrs)) chrs[$1]=1
#  if (! (cm in cms)) cms[cm]=1
#  printf "%5i %5i  %.7f  %0.7f\n",ibdid1,ibdid2,$6*2,$7 >> prefix""$1"."cm
}
```

gzip mibd files

```
END {
  for (m in mcm){
    cm = int(mcm[m])
```

```
    # print missing diagonal elements
    for (i in ibdid) {
      if (! (i in done)) {
        printf "%5i %5i  %.7f  %0.7f\n",ibdid[i],ibdid[i],1.0,1.0 >>
prefix""chr"."cm
      }
    }
    system("gzip -f "prefix""chr"."cm)
  }
}
```

D.5.2 Preparation of map file for MIBD generation I - removal of SNPs not used for IBD estimation

Remove SNPs which were excluded from the IBDLD calculations

This info (241 SNPs) extracted from the log file generated by IBDLD and saved in 1 column (1 SNP per line).

R script used to join extracted SNPs on to map file used for IBDLD program

```
library(dplyr)
```

Read in IBDLD map file (full list of variants)

```
IBDLDmap <- read.csv(file = "IBDLD_map_file.map.txt",header = FALSE,
                             sep = "\t")
```

Read in filtered SNP file (filtered out by IBDLD, want to delete from map file for MIBD)

```
filteredSNPs <- read.csv(file = "filteredSNPs_IBDLD_99%.txt", header =
FALSE)
```

Add in a column to filtered file to indicate which are the SNPs to remove

```
filteredSNPs$V2 <- "to_remove"
```

Name V2 column in files to use for joining

```
colnames(IBDLDmap)[2] <- "chr_bp"
colnames(filteredSNPs)[1] <- "chr_bp"
```

Join full IBDLD file with filtered SNPs file

```
fullMAP <- full_join(IBDLDmap, filteredSNPs, by = "chr_bp")
```

write the output as a text file

```
write.table(fullMAP, file = "IBDLD_map_with_SNPs_to_remove.txt",
      sep = "\t",
      quote = FALSE,
      row.names=FALSE,
      col.names=FALSE,
      na = "")
```

Open file in Excel and sort based on “to remove” column. Delete identified rows.

D.5.3 Preparation of map file for MIBD generation II - selection of 1SNP per cM markers

Switch cM and bp columns in excel, so new map file reads chr, chr_bp, bp,cM

Use INT function in excel to create a new column which rounds cM position down to the nearest integer.

Then use “remove duplicates” while selecting the chromosome and INT columns to just retain 1 marker per cM (retains the first value in the list and removes all duplicates below).

Remove INT column, to retain the 4 original columns.

Subset the full map file into files for each chromosome

D.5.4 MIBD file generation (loop command)

loop command for using the custom awk script with individual chromosome map files, IBDLD output files and pedigree indexing files to generate MIBD files

```
for chr in {1..22}; do \
zcat file_path_to/POAGdate_${chr}.ibdtxt.gz | awk \
-f ibdld2mibd.awk -v prefix=MIBD/mibd. -v indexout=pedindex.out \
-v indexcde=pedindex.cde -v map=chr${chr}MAPforMIBD.txt & done
```

D.5.5 MIBD file processing

grm program written by Juan Peralta, University of Texas Rio Grande

`-z` truncates negative values to zero

`-m` maximum number of iterations allowed for convergence

`-t` tolerance parameter removed (default is $1x10^{-6}$)

```
for f in mibd.*; do \
LD_LIBRARY_PATH=. ./grm nearest_psd -z -m 400 -i linearsymmetric -o linear
<(zcat $f | awk 'NR==1 {print "id1,id2,matrix1"} {print $1","$2","$3}') |
awk -v mibd=$f 'BEGIN {FS=" ";cmd="zcat "mibd; while(cmd | getline){
d7[$1,$2]=$4};FS=","} NR>1 {printf "%5i %5i  %.7f  %0.7f\n",
$1,$2,$3,d7[$1,$2]}' | gzip -c > processed_$f; done
```

D.6 Empirical kinship file (phi2) preparation

The `prefix_genome.kinship` IBDLD output file was used as the starting point for the phi2 file.

The 15th column was multiplied by 2 (in a new column). The ID numbers and new phi2 column were retained and all other columns deleted.

Headers of id1, id2 and marker1 for the 3 retained columns.

File saved as a .csv (for the correct formatting) and then the extenstion was deleted.

New phi2 file zipped and moved to the folder where the linkage analysis is conducted.

D.7 Ascertainment correction tcl script for IOP and IOPmed

Similar modelling commands as in Appendix A.8, except polymod and maximise are used instead of polygenic -screen

This script shows constraint of MaxIOP mean, SD and beta values for the covariates

```
proc ACscript_TRAIT {} {
      model new
      trait IOP
      covariate age^1,2#sex
      polymod
      constrain mean = 15.42639902
      parameter mean = 15.42639902
      constrain sd = 3.338103473
      parameter sd = 3.338103473
      constrain bage = 0.01620147599
      parameter bage = 0.01620147599
      constrain bsex = -0.164419209
      parameter bsex = -0.164419209
      constrain <bage*sex> = 0.00680895065
      parameter bage*sex = 0.00680895065
      constrain <bage^2> = -0.0008252308106
      parameter bage^2 = -0.0008252308106
      constrain <bage^2*sex> = -0.0004636277018
      parameter bage^2*sex = -0.0004636277018
      matrix load -sample V25_phi2.gz phi2
      option MergeAllPeds 1
      maximize
      model save IOP/null0
      model
      }
```

D.8 Ascertainment correction tcl script for maximum VCDR

Similar modelling commands as in Appendix A.8, except polymod and maximise are used instead of polygenic -screen

This script shows constraint of MaxVCDR mean, SD and beta values for the covariates

```
proc ACscript_VCDR {} {
      model new
      trait VCDR
      covariate age^1,2#sex
      polymod
      constrain mean = 4.46177549
      parameter mean = 4.46177549
      constrain sd = 2.020809901
      parameter sd = 2.020809901
      constrain bsex = -0.355021133
      parameter bsex = -0.355021133
      constrain bage = 0.02753255984
      parameter bage = 0.02753255984
      constrain <bage*sex> = -0.004957120049
      parameter bage*sex = -0.004957120049
      constrain <bage^2> = -0.0007189157873
      parameter bage^2 = -0.0007189157873
      constrain <bage^2*sex> = 0.0002952706829
      parameter bage^2*sex = 0.0002952706829
      matrix load -sample V25_phi2.gz phi2
      option MergeAllPeds 1
      maximize
      model save VCDR/null0
      model
      }
```

D.9 Linkage analysis with ascertainment correction commands for SOLAR

load pedigree and phenotype files

```
load ped pedigree_file.ped
load phen phenotype_file.phen

model new
trait TRAIT
```

load tcl script which includes covariates, constraints, loading phi2 and MergeAllPeds

no need to type .tcl at the end

for a tcl script called ACscript_TRAIT.tcl

```
ACscript_TRAIT
```

provide location of MIBD files

```
mibddir file_path_to/MIBD
mibddir
```

run linkage for all the chromosomes

```
chromosome 1-22    # or chromosome all
```

interval of 1 marker

```
interval 1
```

run the multipoint linkage analysis for QTLs with a LOD ≥ 3 and do additional passes until the largest peak does not reach LOD =2

```
multipoint 3 2
```

Appendix E MGA results for QTL analysis

The following table contains the measured genotype association results for the IOPmed trait with $p \leq 0.05$ from within the linkage intervals described in Chapter 5.3.1, broken down into MAC present within each family.

chr_bp = chromosome number and base pair position (hg19)
beta-value = effect size (mmHg), non-transformed data
Included in MGA (MAC) = minor allele count of genotype data with accompanying phenotype data included in the measured genotype analysis
Total Genotyped (MAC) = total minor allele count genotyped in the families
Rows in bold = a single family contributing ≥ 90% of the MAC and tested for linkage conditional on the measured genotype

chr_bp	p-value	beta-value	Included in MGA (MAC)	Family 93001	Family 95002	Family 98002	Family GTAS04	Family GTAS54	Total Genotyped (MAC)
chr2_148687731	0.03672	-3.10403	71	15	3	33	15	6	72
chr2_149881029	0.021112	-0.31372	72	23	13	20	5	10	71
chr2_150327599	0.009312	2.32815	11	1	0	2	0	8	11
chr2_152220510	0.012683	-5.86676	9	0	0	1	7	2	10
chr2_152226616	0.012683	-5.86675	9	0	0	1	7	2	10
chr2_152300147	**0.00534**	**-9.36015**	4	**0**	**0**	**0**	**5**	**0**	**5**
chr2_152317669	**0.021508**	**1.71079**	34	**0**	**0**	**36**	**0**	**0**	**36**
chr2_152321091	**0.032912**	**-6.79727**	5	**0**	**0**	**5**	**0**	**0**	**5**
chr2_152331769	0.009528	0.95387	60	5	2	43	9	4	63
chr2_152331995	**0.021508**	**1.71079**	34	**0**	**0**	**36**	**0**	**0**	**36**
chr2_152387592	**0.051654**	**2.32988**	16	**16**	**0**	**0**	**0**	**0**	**16**
chr2_152658296	0.015969	3.07686	5	1	2	2	0	0	5
chr2_152695190	0.018201	3.13282	6	0	0	1	3	2	6
chr2_152976670	0.039887	-4.96469	10	3	1	5	0	1	10
chr2_152976702	0.039887	-4.96469	10	3	1	5	0	1	10
chr2_153032341	0.013664	2.03996	13	0	0	0	3	10	13
chr2_154334499	0.047524	-0.66388	250	71	32	87	39	25	254
chr2_155555878	0.029176	1.44245	15	2	1	0	0	12	15
chr2_157441670	0.014453	1.74038	17	6	0	0	0	11	17
chr2_157442354	**0.033259**	**-6.79592**	5	**0**	**0**	**5**	**0**	**0**	**5**
chr2_158157361	0.011068	1.42136	464	111	71	179	65	46	472
chr2_158177675	**0.001213**	**7.99404**	6	**0**	**6**	**0**	**0**	**0**	**6**
chr2_158385315	**0.047841**	**2.41229**	15	**15**	**0**	**0**	**0**	**0**	**15**
chr2_159166069	0.016792	1.70981	471	112	66	175	74	52	479
chr2_159517926	0.018676	-7.17334	8	0	2	5	1	0	8
chr2_159536990	**0.001276**	**4.61379**	8	**0**	**0**	**0**	**0**	**8**	**8**

chr_bp	p-value	beta-value	Included in MGA (MAC)	Family 93001	Family 95002	Family 98002	Family GTAS04	Family GTAS54	Total Genotyped (MAC)
chr2_159537297	0.034126	2.44277	7	1	1	0	5	0	7
chr2_159954175	0.018085	-2.88489	115	29	16	25	19	27	116
chr2_160569020	**0.044099**	**1.57866**	24	**23**	**0**	**0**	**0**	**1**	**24**
chr2_160604379	**0.044099**	**1.57866**	24	**23**	**0**	**0**	**0**	**1**	**24**
chr2_160604768	**0.044099**	**1.57866**	24	**23**	**0**	**0**	**0**	**1**	**24**
chr2_160604812	0.030669	-2.77613	160	33	18	62	33	16	162
chr2_160625727	0.001138	-3.52924	149	35	16	53	33	14	151
chr2_160627317	0.005055	-3.06511	97	27	12	35	17	8	99
chr2_160628155	0.000701	-3.57268	147	35	17	52	32	13	149
chr2_160639970	0.000666	-3.64955	145	32	16	52	33	14	147
chr2_160660189	**0.028976**	**1.62844**	25	**23**	**1**	**0**	**0**	**1**	**25**
chr2_160661107	**0.048186**	**1.53375**	24	**23**	**0**	**0**	**0**	**1**	**24**
chr2_160737622	**0.005487**	**-9.35673**	4	**0**	**0**	**0**	**5**	**0**	**5**
chr2_160919050	**0.037884**	**3.37375**	5	**0**	**0**	**0**	**5**	**0**	**5**
chr2_161349966	**0.041938**	**2.16138**	17	**17**	**0**	**0**	**0**	**0**	**17**
chr2_163128824	0.026782	-3.06399	344	94	37	120	53	46	350
chr2_165946416	**0.044685**	**0.86674**	14	**1**	**0**	**0**	**13**	**0**	**14**
chr2_166032775	0.012	-3.67418	39	3	7	22	1	6	39
chr2_166740469	0.033389	0.95825	32	23	5	2	0	2	32
chr2_166810232	**0.007935**	**3.42906**	9	**0**	**0**	**0**	**0**	**9**	**9**
chr2_167144974	0.000885	-3.44397	198	46	16	88	32	22	204
chr2_167144995	0.002962	-3.35093	171	44	16	61	32	21	174
chr2_167145142	0.003462	-3.26942	162	43	15	59	30	18	165
chr2_167313451	0.033368	-3.29698	84	20	10	32	18	5	85
chr2_167321993	**0.020364**	**2.30886**	20	**20**	**0**	**0**	**0**	**0**	**20**
chr2_168102774	0.017428	-4.24146	28	7	12	3	2	4	28
chr2_168104183	0.033646	-6.64238	5	1	0	4	0	0	5
chr2_168726592	**0.047905**	**2.64191**	8	**0**	**0**	**9**	**0**	**0**	**9**
chr2_168726659	0.00255	0.50725	102	36	7	35	15	10	103
chr2_169952233	0.02097	-3.10096	85	16	5	29	22	13	85
chr2_170361334	0.005038	-3.92889	56	12	12	22	6	4	56
chr2_170367099	0.005038	-3.92889	56	12	12	22	6	4	56
chr2_170403106	0.020682	0.60382	43	8	15	6	1	13	43
chr2_170430388	0.026969	-0.25604	113	20	21	34	26	15	116
chr2_170597913	**0.053042**	**-7.18340**	8	**0**	**0**	**9**	**0**	**0**	**9**
chr2_170762566	0.053739	-0.42084	80	15	18	15	19	13	80
chr2_171242761	0.045859	-1.64938	82	19	2	39	18	6	84
chr2_171570151	0.039398	-1.41272	139	13	18	48	28	35	142
chr2_171570488	0.001828	-1.79131	316	71	25	140	47	39	322

chr_bp	p-value	beta-value	Included in MGA (MAC)	Family 93001	Family 95002	Family 98002	Family GTAS04	Family GTAS54	Total Genotyped (MAC)
chr2_171573597	**0.022572**	**4.43855**	16	**16**	**0**	**0**	**0**	**0**	**16**
chr2_171574217	0.002474	-2.06115	343	74	26	155	50	44	349
chr2_171849612	0.010833	2.07926	112	26	30	26	19	13	114
chr2_171910313	0.010833	2.07926	112	26	30	26	19	13	114
chr2_172174150	0.02009	-1.34278	343	88	46	109	62	45	350
chr2_172174151	0.02009	-1.34278	343	88	46	109	62	45	350
chr2_172174263	0.021883	-1.41399	351	89	48	112	63	46	358
chr2_172175068	0.026575	-1.37750	337	84	40	112	62	46	344
chr2_172188368	0.039098	1.26829	136	25	24	67	11	10	137
chr2_172412835	0.047152	-1.36000	104	26	18	34	21	9	108
chr2_172412836	0.047152	-1.36000	104	26	18	34	21	9	108
chr2_172585298	0.030574	1.89826	41	7	3	23	3	7	43
chr2_172650165	0.026833	-3.16920	19	2	0	9	6	2	19
chr2_172864845	0.017563	-0.99053	264	63	31	94	37	44	269
chr2_172953525	**0.007935**	**5.02290**	9	**0**	**0**	**0**	**0**	**9**	**9**
chr2_173352103	0.040329	-1.51678	82	9	11	43	12	10	85
chr2_173369000	0.011837	-1.91599	76	9	9	43	11	7	79
chr2_173461090	0.053921	2.39005	21	1	0	9	2	9	21
chr2_173463649	**0.022572**	**4.43855**	16	**16**	**0**	**0**	**0**	**0**	**16**
chr2_173588663	0.029383	2.20717	20	1	2	6	0	11	20
chr2_173686569	0.024137	-8.81417	3	69	33	106	23	23	254
chr3_71731348	0.033903	1.80156	56	3	15	10	9	19	56
chr3_73024350	0.002121	1.63430	313	90	40	110	31	46	317
chr3_73437185	0.011651	4.46663	10	0	0	1	0	9	10
chr3_75471002	**0.018981**	**4.23992**	10	**0**	**0**	**0**	**0**	**10**	**10**
chr3_75471289	0.024081	3.35708	13	4	9	0	0	0	13
chr3_75471644	0.041758	2.70174	14	5	9	0	0	0	14
chr3_75471832	0.041758	2.70174	14	5	9	0	0	0	14
chr3_75476615	0.041758	2.70174	14	5	9	0	0	0	14
chr3_75477000	0.041758	2.70174	14	5	9	0	0	0	14
chr3_77089699	0.040147	1.31813	68	23	14	14	5	13	69
chr6_143091769	0.010803	1.94990	75	14	29	17	8	7	75
chr6_143780313	**0.003533**	**6.06502**	8	**0**	**8**	**0**	**0**	**0**	**8**
chr6_143876124	0.046952	-1.18018	114	29	19	33	12	21	114
chr6_143890039	0.002539	-1.58743	340	74	47	119	62	42	344
chr6_143890049	0.00483	-1.45827	341	75	47	119	62	42	345
chr6_143890097	0.006376	-1.43845	342	75	48	119	62	42	346
chr6_144150609	0.033318	1.81971	55	12	12	20	7	5	56
chr6_144150831	0.03243	1.80093	56	12	12	21	7	5	57

chr_bp	p-value	beta-value	Included in MGA (MAC)	Family 93001	Family 95002	Family 98002	Family GTAS04	Family GTAS54	Total Genotyped (MAC)
chr6_147887092	0.003194	6.02032	11	0	9	1	1	0	11
chr6_147887218	0.030777	1.11850	310	53	53	129	44	37	316
chr6_147887569	0.019099	1.23319	161	37	34	46	28	18	163
chr6_147887890	0.002295	3.70064	24	3	16	1	2	2	24
chr6_147888022	0.031363	1.09751	316	53	55	132	45	37	322
chr6_147890370	0.000369	4.26638	23	3	16	1	2	1	23
chr6_150239330	0.034563	-2.01054	69	7	6	33	5	19	70
chr6_150240829	0.034563	-2.01054	69	7	6	33	5	19	70
chr6_150570837	0.031942	-3.24875	13	3	0	5	3	2	13
chr6_150570867	0.031942	-3.24876	13	3	0	5	3	2	13
chr6_150570940	0.031942	-3.24876	13	3	0	5	3	2	13
chr6_150719710	0.016934	-1.49973	273	79	18	95	54	32	278
chr6_151152230	**0.034745**	**4.45393**	8	**0**	**8**	**0**	**0**	**0**	**8**
chr6_151670690	0.026239	-2.64579	18	1	3	11	3	0	18
chr6_151672285	0.039019	-2.42027	19	1	4	11	3	0	19
chr6_151674270	0.026239	-2.64579	18	1	3	11	3	0	18
chr6_151674326	0.021986	1.52424	332	81	63	108	49	37	338
chr6_151678315	0.047685	1.33775	333	85	61	109	48	36	339
chr6_151678385	0.044608	1.31602	331	83	62	109	46	37	337
chr6_151678499	0.036508	1.43204	338	85	63	110	49	37	344
chr6_151679198	0.036508	1.43204	338	85	63	110	49	37	344
chr6_151685528	0.011532	1.61660	311	81	61	103	45	27	317
chr6_151686905	0.012438	1.58705	312	82	61	103	45	27	318
chr6_151726300	0.044336	-1.50777	94	17	2	47	13	16	95
chr6_151766552	0.021243	-2.49787	41	8	4	26	3	1	42
chr6_151773504	0.045027	-1.24940	127	28	7	56	13	24	128
chr6_151789912	0.035681	-2.25948	42	8	4	27	3	1	43
chr7_55238268	0.021056	1.21829	146	56	11	29	38	12	146
chr7_55238464	0.034338	1.15142	144	54	11	28	38	13	144
chr7_55268916	0.040816	-2.09606	38	7	4	12	7	9	39
chr7_55539538	**0.011415**	**4.51296**	11	**0**	**0**	**12**	**0**	**0**	**12**
chr7_55540132	0.004189	-1.89892	89	20	8	48	10	5	91
chr7_55714282	0.015883	-1.34902	269	68	23	99	60	25	275
chr7_55749559	0.007659	-2.02621	66	11	6	39	6	7	69
chr7_55749603	0.017199	2.07883	120	30	21	46	13	11	121
chr7_55862375	0.051692	1.67622	29	4	6	6	6	8	30
chr7_56087300	0.027421	1.65745	76	29	5	23	7	13	77
chr7_56087319	0.017133	1.76859	77	29	6	23	7	13	78
chr7_56087364	0.017133	1.76859	77	29	6	23	7	13	78

chr_bp	p-value	beta-value	Included in MGA (MAC)	Family 93001	Family 95002	Family 98002	Family GTAS04	Family GTAS54	Total Genotyped (MAC)
chr7_56087365	0.017133	1.76859	77	29	6	23	7	13	78
chr7_56087374	0.017133	1.76859	77	29	6	23	7	13	78
chr7_56087379	0.017133	1.76859	77	29	6	23	7	13	78
chr7_56087399	0.013163	1.87845	76	29	6	22	7	13	77
chr7_56087409	0.013163	1.87845	76	29	6	22	7	13	77
chr7_56087423	0.026111	1.70278	74	28	6	21	7	13	75
chr7_56088811	0.024994	1.76331	72	28	5	20	7	13	73
chr7_56088825	0.022266	1.79305	73	28	5	21	7	13	74
chr7_56088902	0.007411	2.02720	72	29	5	21	7	11	73
chr7_56088907	0.007411	2.02720	72	29	5	21	7	11	73
chr7_56088908	0.007411	2.02720	72	29	5	21	7	11	73
chr7_57528997	0.046438	0.88542	180	22	28	65	44	23	182
chr7_57529392	0.046438	0.88542	180	22	28	65	44	23	182
chr7_57530627	0.046438	0.88542	180	22	28	65	44	23	182
chr7_57530731	0.043	0.91429	176	22	28	64	42	22	178
chr7_57530733	0.043	0.91429	176	22	28	64	42	22	178
chr7_57530786	0.033895	0.93837	219	25	29	88	47	32	221
chr7_57531189	0.045244	0.86260	203	25	26	77	46	31	205
chr7_57531524	0.049125	0.88860	179	22	28	65	43	23	181
chr7_57531737	0.046438	0.88542	180	22	28	65	44	23	182
chr7_57531922	0.048092	0.87710	181	22	28	66	44	23	183
chr7_62809629	**0.01444**	**5.99714**	8	**8**	**0**	**0**	**0**	**0**	**8**
chr7_63674464	0.006821	3.97540	22	0	8	15	0	0	23
chr7_64254791	0.011112	1.26726	326	77	54	124	46	30	331
chr7_64254901	0.008929	1.39304	333	80	56	124	47	31	338
chr7_64292695	0.010306	-1.40817	158	36	16	58	27	25	162
chr7_64293217	0.009827	-1.41627	158	36	17	56	27	25	161
chr7_64293575	0.004743	-1.55064	159	36	17	57	27	25	162
chr7_64293731	0.008494	-1.43742	160	36	17	58	27	25	163
chr7_64311698	0.045154	3.80551	10	1	2	7	0	0	10
chr7_64312825	0.045154	3.80551	10	1	2	7	0	0	10
chr7_64313784	0.031492	-2.23034	453	96	70	170	73	52	461
chr7_64389067	0.045154	3.80551	10	1	2	7	0	0	10
chr7_64389478	0.045154	3.80551	10	1	2	7	0	0	10
chr7_64437820	0.021674	-2.43071	459	98	72	172	73	52	467
chr7_64437832	0.021674	-2.43071	459	98	72	172	73	52	467
chr7_64661070	0.035846	1.72419	202	45	31	70	33	27	206
chr7_65159983	0.041272	1.03087	203	44	36	64	36	28	208
chr7_65425894	0.027299	-1.14825	245	67	26	97	35	23	248

chr_bp	p-value	beta-value	Included in MGA (MAC)	Family 93001	Family 95002	Family 98002	Family GTAS04	Family GTAS54	Total Genotyped (MAC)
chr7_65579838	0.011534	-4.49428	7	1	2	3	1	0	7
chr7_65825257	0.013809	-1.86826	51	9	5	22	10	7	53
chr7_66019649	0.023312	-1.72741	50	8	5	22	10	7	52
chr7_66038388	0.006026	1.95280	67	12	13	28	12	3	68
chr7_66098384	0.023312	-1.72741	50	8	5	22	10	7	52
chr7_66107971	0.030377	-1.68291	51	8	5	22	11	7	53
chr7_66297377	0.031411	-1.66180	52	8	5	23	11	7	54
chr7_66421143	0.023312	-1.72741	50	8	5	22	10	7	52
chr7_66760636	0.010972	1.27490	151	48	23	62	8	12	153
chr7_66762277	0.024745	1.03540	205	62	27	83	12	23	207
chr7_66764392	0.011411	-2.63992	36	7	4	16	8	1	36
chr7_66767623	0.044782	1.28099	116	22	24	50	8	14	118
chr7_66785137	0.025835	1.33131	188	60	27	75	11	18	191
chr7_66786212	0.002585	-3.00772	36	5	5	19	5	2	36
chr7_66786384	0.04138	1.28658	186	60	27	73	11	18	189
chr7_71135043	**0.01308**	**-3.92959**	10	**0**	**0**	**10**	**0**	**0**	**10**
chr7_71178166	0.000971	-3.26140	25	2	4	14	4	2	26
chr7_71249529	**0.014784**	**4.27507**	13	**0**	**0**	**14**	**0**	**0**	**14**
chr7_71249770	**0.014784**	**4.27507**	13	**0**	**0**	**14**	**0**	**0**	**14**
chr7_72333411	**0.01444**	**5.99714**	8	**8**	**0**	**0**	**0**	**0**	**8**
chr7_72334664	0.033018	-2.43115	21	8	0	10	2	1	21
chr7_72338657	0.002418	2.62723	64	22	12	15	9	7	65
chr7_72338825	0.040606	-2.42225	21	9	0	9	2	1	21
chr7_72350159	0.025911	1.11899	288	87	54	89	40	22	292
chr7_72350361	0.033798	1.11449	301	87	55	92	43	28	305
chr7_72417063	0.015279	1.30693	266	75	48	82	42	23	270
chr7_72418618	0.025961	2.33047	46	10	5	15	13	3	46
chr7_72469999	0.052846	1.76094	140	33	27	57	7	17	141
chr7_72476293	0.000313	6.59143	15	11	4	0	0	0	15
chr7_72850305	0.001494	5.65415	14	10	4	0	0	0	14
chr7_72855568	**0.00498**	**8.45610**	7	**0**	**0**	**8**	**0**	**0**	**8**
chr7_72855662	0.029575	3.38497	12	3	2	4	3	0	12
chr7_73020301	0.014512	-5.20449	479	112	72	173	73	56	486
chr7_73096993	0.015851	-1.11394	159	36	29	62	16	19	162
chr7_73097654	0.010458	-1.21946	158	36	29	61	16	19	161
chr7_73536579	**0.005095**	**8.43094**	7	**0**	**0**	**8**	**0**	**0**	**8**
chr7_73814702	0.038331	-3.63691	15	1	0	11	1	2	15
chr7_74193642	0.039346	2.03234	12	2	1	8	1	0	12
chr7_75046785	0.043941	1.45864	159	39	28	74	7	14	162

chr_bp	p-value	beta-value	Included in MGA (MAC)	Family 93001	Family 95002	Family 98002	Family GTAS04	Family GTAS54	Total Genotyped (MAC)
chr7_75048172	0.004798	-3.94712	8	0	1	6	1	0	8
chr7_75051375	0.001166	1.69773	260	56	47	111	30	20	264
chr7_75115535	0.03207	2.81049	20	0	0	17	3	0	20
chr7_75141695	**0.014784**	**4.27507**	13	**0**	**0**	**14**	**0**	**0**	**14**
chr7_75167225	0.046961	1.43214	136	34	23	65	5	12	139
chr7_75544455	0.038538	-1.41599	172	15	40	59	39	21	174
chr7_76038848	0.00086	4.17345	33	12	0	20	2	0	34
chr7_76111938	0.006738	4.24846	25	12	0	13	1	0	26
chr7_76239511	0.038826	-4.33925	6	0	0	5	2	0	7
chr7_76610344	**0.016849**	**-4.59266**	8	**0**	**0**	**8**	**0**	**0**	**8**
chr7_76619625	0.030922	-1.78738	76	21	7	19	11	19	77
chr7_76629626	0.03436	-1.69654	158	44	15	60	16	25	160
chr7_76631549	0.033471	-1.24804	114	27	10	42	22	16	117
chr7_76669334	0.007761	-1.54339	112	26	11	40	23	16	116
chr7_76681663	0.022883	-1.31064	96	20	11	33	24	10	98
chr7_76681764	0.000241	-2.34286	101	22	6	39	22	15	104
chr7_76682235	0.008884	1.86510	102	33	22	26	15	8	104
chr7_77167371	0.0149	-3.34340	14	0	0	4	5	5	14
chr7_77553078	0.041947	-2.72359	15	1	2	4	3	5	15
chr9_137779026	0.041546	1.16452	171	20	36	82	17	19	174
chr9_138012886	0.019012	-1.33445	140	24	40	39	28	12	143
chr9_138453641	0.018304	2.69157	41	16	11	4	1	9	41
chr9_138515632	0.020238	-4.13968	14	3	0	10	3	0	16
chr9_138515991	0.041242	-1.89073	60	24	8	24	4	2	62
chr9_138585419	**0.010369**	**7.34035**	7	**0**	**0**	**7**	**0**	**0**	**7**
chr9_138586966	0.047378	2.35557	451	103	71	162	71	51	458
chr9_138586967	0.047378	2.35557	451	103	71	162	71	51	458
chr9_138700604	**0.036658**	**7.00902**	6	**0**	**0**	**6**	**0**	**0**	**6**
chr9_138715799	0.026967	1.95662	31	7	3	10	11	1	32
chr9_138836941	0.007331	1.52518	147	40	21	43	23	21	148
chr9_138908250	0.027168	-1.87137	55	7	5	24	17	2	55
chr9_139235606	0.010488	2.60756	35	21	1	4	3	7	36
chr9_139256468	0.01519	2.24644	38	21	1	5	6	6	39
chr9_139256541	0.034995	2.08737	36	21	0	5	5	6	37
chr9_139296703	**0.006599**	**5.18076**	15	**15**	**0**	**0**	**0**	**0**	**15**
chr9_139301708	0.000686	4.67169	22	15	0	0	1	6	22
chr9_139351941	0.00505	3.62115	23	16	0	0	1	6	23
chr9_139566823	0.017659	-2.29652	443	92	68	174	66	51	**451**
chr9_139649934	0.03615	2.90053	15	11	4	0	0	0	15

chr_bp	p-value	beta-value	Included in MGA (MAC)	Family 93001	Family 95002	Family 98002	Family GTAS04	Family GTAS54	Total Genotyped (MAC)
chr9_139726239	0.016884	1.60980	373	97	62	117	56	46	378
chr9_139964447	0.051812	-1.28576	133	27	8	69	26	5	135
chr9_139975195	0.003499	3.17817	38	2	28	2	4	2	38
chr9_140002995	0.018575	2.57932	42	21	2	1	3	15	42
chr9_140003529	0.000044	3.19434	78	20	32	2	7	17	78
chr9_140112447	0.018819	1.72059	70	24	5	13	8	20	70
chr9_140194793	0.047551	-3.90527	5	1	0	2	2	0	5
chr9_140277794	0.049938	-1.96111	35	8	3	13	8	5	37
chr9_140342693	**0.006599**	**5.18075**	15	**15**	**0**	**0**	**0**	**0**	**15**
chr9_140357943	0.001956	2.84684	46	17	7	7	2	13	46
chr9_140400464	0.04592	-3.70320	7	0	5	0	2	0	7
chr15_20777925	0.018561	1.40576	154	36	29	56	21	16	158
chr15_20778593	0.016736	1.71828	36	14	10	5	2	5	36
chr15_21932917	0.002253	2.34371	145	38	23	38	25	24	148
chr15_21936054	0.002639	2.30104	145	39	23	38	24	24	148
chr15_21936801	0.018173	2.04138	60	9	13	21	4	14	61
chr15_21936861	0.047411	1.67117	135	35	23	33	23	24	138
chr15_21937528	0.004425	2.19615	145	39	23	37	25	24	148
chr15_21937536	0.002367	2.33134	146	39	23	38	25	24	149
chr15_22014916	**0.019128**	**3.41153**	14	**1**	**0**	**0**	**0**	**13**	**14**
chr15_22015598	0.028673	-0.89489	230	71	22	82	35	23	233
chr15_22015940	0.028576	-1.06769	280	73	32	105	46	28	284
chr15_22016334	0.01266	-1.73392	82	33	12	29	8	2	84
chr15_22145400	0.049186	-1.11887	193	7	3	29	11	6	56
chr15_22382655	0.033109	1.49971	142	56	27	101	44	34	262
chr15_22383066	0.050809	1.08748	307	37	40	32	11	15	135
chr15_22835939	**0.017002**	**4.09186**	11	**0**	**0**	**0**	**0**	**11**	**11**
chr15_22863037	0.028362	-3.43487	9	0	1	8	0	0	9
chr15_23043644	**0.017002**	**4.09186**	11	**0**	**0**	**0**	**0**	**11**	**11**
chr15_23046770	0.011603	1.55032	338	66	58	134	48	37	343
chr15_23052632	0.002385	1.86423	349	66	63	139	48	38	354
chr15_23191502	0.035843	-1.16060	246	61	44	84	37	23	249
chr15_23200323	0.012758	-1.47988	183	34	30	76	30	16	186
chr15_23201566	0.012426	1.30457	276	63	41	97	42	37	280
chr15_23377447	**0.008723**	**-5.39482**	5	**0**	**0**	**5**	**0**	**0**	**5**
chr15_23377450	**0.008723**	**-5.39481**	5	**0**	**0**	**5**	**0**	**0**	**5**
chr15_23445715	0.0154	2.23850	49	12	10	16	4	9	51
chr15_23890600	**0.034234**	**6.11886**	6	**6**	**0**	**0**	**0**	**0**	**6**
chr15_24410011	0.003292	2.42020	62	5	13	23	14	8	63

chr_bp	p-value	beta-value	Included in MGA (MAC)	Family 93001	Family 95002	Family 98002	Family GTAS04	Family GTAS54	Total Genotyped (MAC)
chr15_24410100	0.003292	2.42020	62	5	13	23	14	8	63
chr15_24410204	0.007524	2.24996	60	5	12	23	13	8	61
chr15_24412931	0.003016	3.15175	38	3	11	14	6	5	39
chr15_24414271	0.003292	2.42020	62	5	13	23	14	8	63
chr15_24924663	0.048426	-1.03352	106	14	17	43	22	10	106
chr15_25101838	0.040159	1.57999	51	6	3	32	2	9	52
chr15_25227494	0.045391	-1.01635	140	33	22	53	26	8	142
chr15_25228550	0.023891	1.88528	203	51	32	68	33	23	207
chr15_25362564	0.035889	1.64680	79	8	20	39	9	6	82
chr15_25380969	0.010376	-5.84234	6	0	0	4	2	0	6
chr15_25438493	0.010041	1.54228	107	11	3	59	14	21	108
chr15_25925094	0.014669	-1.27939	209	54	25	68	40	26	213
chr15_25953198	0.010223	-3.49603	15	0	3	5	3	4	15
chr15_26790609	**0.048712**	**3.09960**	10	**0**	**0**	**0**	**0**	**10**	**10**
chr15_28259941	**0.048712**	**3.09960**	10	**0**	**0**	**0**	**0**	**10**	**10**
chr15_28474713	**0.000491**	**10.34104**	5	**0**	**5**	**0**	**0**	**0**	**5**
chr15_29992925	0.008521	1.27905	258	67	34	88	28	27	244

References

1. Tham YC, Li X, Wong TY, et al. Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. Ophthalmology. 2014;121(11):2081-90.

2. Quigley HA. Glaucoma. Lancet. 2011;377(9774):1367-77.

3. Weinreb RN, Aung T, Medeiros FA. The pathophysiology and treatment of glaucoma: a review. JAMA. 2014;311(18):1901-11.

4. deLuise VP, Anderson DR. Primary infantile glaucoma (congenital glaucoma). Survey of ophthalmology. 1983;28(1):1-19.

5. Kwon YH, Fingert JH, Kuehn MH, et al. Primary open-angle glaucoma. The New England journal of medicine. 2009;360(11):1113-24.

6. Quigley HA, Broman AT. The number of people with glaucoma worldwide in 2010 and 2020. Br J Ophthalmol. 2006;90(3):262-7.

7. Kapetanakis VV, Chan MP, Foster PJ, et al. Global variations and time trends in the prevalence of primary open angle glaucoma (POAG): a systematic review and meta-analysis. Br J Ophthalmol. 2016;100(1):86-93.

8. Keel S, Xie J, Foreman J, et al. Prevalence of glaucoma in the Australian National Eye Health Survey. Br J Ophthalmol. 2019;103(2):191-5.

9. Tunnel Vision. The economic impact of Primary Open Angle Glaucoma - A Dynamic Economic Model. 2008.

10. Wójcik-Gryciuk A, Skup M, Waleszczyk WJ. Glaucoma–state of the art and perspectives on treatment. Restorative neurology and neuroscience. 2015;34(1):107-23.

11. Weinreb RN, Khaw PT. Primary open-angle glaucoma. Lancet. 2004;363(9422):1711-20.

12. Bailey JNC, Loomis SJ, Kang JH, et al. Genome-wide association analysis identifies TXNRD2, ATXN2 and FOXC1 as susceptibility loci for primary open-angle glaucoma. Nature Genetics. 2016;48(2):189-94.

13. Janssen SF, Gorgels TG, Ramdas WD, et al. The vast complexity of primary open angle glaucoma: disease genes, risks, molecular mechanisms and pathobiology. Prog Retin Eye Res. 2013;37:31-67.

14. Foster PJ, Buhrmann R, Quigley HA, et al. The definition and classification of glaucoma in prevalence surveys. British journal of ophthalmology. 2002;86(2):238-42.

15. Doucette LP, Rasnitsyn A, Seifi M, et al. The interactions of genes, age, and environment in glaucoma pathogenesis. Survey of ophthalmology. 2015;60(4):310-26.

16. Boland MV, Quigley HA. Risk factors and open-angle glaucoma: classification and application. J Glaucoma. 2007;16(4):406-18.

17. Bonomi L, Marchini G, Marraffa M, et al. Prevalence of glaucoma and intraocular pressure distribution in a defined population: the Egna-Neumarkt Study. Ophthalmology. 1998;105(2):209-15.

18. Fuse N. Genetic bases for glaucoma. Tohoku J Exp Med. 2010;221(1):1-10.

19. Sakurada Y, Mabuchi F. Advances in glaucoma genetics. Progress in brain research. 2015;220:107-26.

20. Suzuki Y, Iwase A, Araie M, et al. Risk factors for open-angle glaucoma in a Japanese population: the Tajimi Study. Ophthalmology. 2006;113(9):1613-7.

21. Siaudvytyte L, Januleviciene I, Daveckaite A, et al. Literature review and meta-analysis of translaminar pressure difference in open-angle glaucoma. Eye. 2015;29(10):1242-50.

22. Hewitt AW, Craig JE, Mackey DA. Complex genetics of complex traits: the case of primary open-angle glaucoma. Clinical & experimental ophthalmology. 2006;34(5):472-84.

23. Broadway DC. Visual field testing for glaucoma - a practical guide. Community eye health. 2012;25(79-80):66-70.

24. Boland MV, Ervin A-M, Friedman DS, et al. Comparative Effectiveness of Treatments for Open-Angle Glaucoma: A Systematic Review for the U.S. Preventive Services Task Force. Annals of Internal Medicine. 2013;158(4):271-9.

25. Collaborative Normal-Tension Glaucoma Study G. Comparison of glaucomatous progression between untreated patients with normal-tension glaucoma and patients with therapeutically reduced intraocular pressures. American Journal of Ophthalmology. 1998;126(4):487-97.

26. Kass MA, Heuer DK, Higginbotham EJ, et al. The Ocular Hypertension Treatment Study: a randomized trial determines that topical ocular hypotensive medication delays or prevents the onset of primary open-angle glaucoma. Archives of ophthalmology. 2002;120(6):701-13.

27. Abu-Amero K, Kondkar AA, Chalam KV. An Updated Review on the Genetics of Primary Open Angle Glaucoma. Int J Mol Sci. 2015;16(12):28886-911.

28. Kumar S, Malik MA, Goswami S, et al. Candidate genes involved in the susceptibility of primary open angle glaucoma. Gene. 2016;577(2):119-31.

29. Dimasi DP, Burdon KP, Hewitt AW, et al. Genetic investigation into the endophenotypic status of central corneal thickness and optic disc parameters in relation to open-angle glaucoma. Am J Ophthalmol. 2012;154(5):833-42 e2.

30. Mitchell P, Smith W, Attebo K, et al. Prevalence of open-angle glaucoma in Australia: the Blue Mountains Eye Study. Ophthalmology. 1996;103(10):1661-9.

31. Goldmann H, Schmidt T. [Applanation tonometry]. Ophthalmologica Journal international d'ophtalmologie International journal of ophthalmology Zeitschrift fur Augenheilkunde. 1957;134(4):221-42.

32. Mackey DA, Hewitt AW. Genome-wide association study success in ophthalmology. Curr Opin Ophthalmol. 2014;25(5):386-93.

33. Brandt JD. Central corneal thickness—tonometry artifact, or something more? Ophthalmology. 2007;114(11):1963-4.

34. Ramdas WD, Wolfs RC, Hofman A, et al. Lifestyle and risk of developing open-angle glaucoma: the Rotterdam study. Arch Ophthalmol. 2011;129(6):767-72.

35. Fan BJ, Leung YF, Wang N, et al. Genetic and environmental risk factors for primary open-angle glaucoma. Chinese medical journal. 2004;117(5):706-10.

36. Wang R, Wiggs JL. Common and rare genetic risk factors for glaucoma. Cold Spring Harbor perspectives in medicine. 2014;4(12).

37. Iglesias AI, Springelkamp H, Ramdas WD, et al. Genes, pathways, and animal models in primary open-angle glaucoma. Eye (Basingstoke). 2015;29(10):1285-98.

38. Wolfs RC, Klaver CC, Ramrattan RS, et al. Genetic risk of primary open-angle glaucoma. Population-based familial aggregation study. Arch Ophthalmol. 1998;116(12):1640-5.

39. Gong G, Kosoko-Lasaki S, Haynatzki G, et al. Inherited, familial and sporadic primary open-angle glaucoma. Journal of the National Medical Association. 2007;99(5):559-63.

40. Cuellar-Partida G, Craig JE, Burdon KP, et al. Assessment of polygenic effects links primary open-angle glaucoma and age-related macular degeneration. Scientific reports. 2016;6:26885.

41. Teikari JM. Genetic factors in open-angle (simple and capsular) glaucoma. A population-based twin study. Acta Ophthalmol (Copenh). 1987;65(6):715-20.

42. Charlesworth J, Kramer PL, Dyer T, et al. The path to open-angle glaucoma gene discovery: endophenotypic status of intraocular pressure, cup-to-disc ratio, and central corneal thickness. Invest Ophthalmol Vis Sci. 2010;51(7):3509-14.

43. Klein BE, Klein R, Lee KE. Heritability of risk factors for primary open-angle glaucoma: the Beaver Dam Eye Study. Invest Ophthalmol Vis Sci. 2004;45(1):59-62.

44. Chang TC, Congdon NG, Wojciechowski R, et al. Determinants and heritability of intraocular pressure and cup-to-disc ratio in a defined older population. Ophthalmology. 2005;112(7):1186-91.

45. Freeman EE, Roy-Gagnon MH, Descovich D, et al. The heritability of glaucoma-related traits corneal hysteresis, central corneal thickness, intraocular pressure, and choroidal blood flow pulsatility. PloS one. 2013;8(1):e55573.

46. Sheffield VC, Stone EM, Alward WL, et al. Genetic linkage of familial open angle glaucoma to chromosome 1q21-q31. Nat Genet. 1993;4(1):47-50.

47. Wiggs J, Lynch S, Ynagi G, et al. A genomewide scan identifies novel early-onset primary open-angle glaucoma loci on 9q22 and 20p12. The American Journal of Human Genetics. 2004;74(6):1314-20.

48. Pang CP, Fan BJ, Canlas O, et al. A genome-wide scan maps a novel juvenile-onset primary open angle glaucoma locus to chromosome 5q. Molecular vision. 2006;12:85-92.

49. Wang DY, Fan BJ, Chua JK, et al. A genome-wide scan maps a novel juvenile-onset primary open-angle glaucoma locus to 15q. Investigative ophthalmology & visual science. 2006;47(12):5315-21.

50. Sarfarazi M, Akarsu NA, Hossain A, et al. Assignment of a locus (GLC3A) for primary congenital glaucoma (Buphthalmos) to 2p21 and evidence for genetic heterogeneity. Genomics. 1995;30(2):171-7.

51. Fingert JH. Primary open-angle glaucoma genes. Eye (Lond). 2011;25(5):587-95.

52. Fan BJ, Wang DY, Lam DS, et al. Gene mapping for primary open angle glaucoma. Clin Biochem. 2006;39(3):249-58.

53. Miller MA, Fingert JH, Bettis DI. Genetics and genetic testing for glaucoma. Curr Opin Ophthalmol. 2017;28(2):133-8.

54. Souzeau E, Tram KH, Witney M, et al. Myocilin Predictive Genetic Testing for Primary Open-Angle Glaucoma Leads to Early Identification of At-Risk Individuals. Ophthalmology. 2017;124(3):303-9.

55. Stone EM, Fingert JH, Alward WL, et al. Identification of a gene that causes primary open angle glaucoma. Science. 1997;275(5300):668-70.

56. Hewitt AW, Mackey DA, Craig JE. Myocilin allele-specific glaucoma phenotype database. Human mutation. 2008;29(2):207-11.

57. Takamoto M, Araie M. Genetics of primary open angle glaucoma. Jpn J Ophthalmol. 2014;58(1):1-15.

58. Allingham RR, Wiggs JL, De La Paz MA, et al. Gln368STOP myocilin mutation in families with late-onset primary open-angle glaucoma. Invest Ophthalmol Vis Sci. 1998;39(12):2288-95.

59. Rezaie T, Child A, Hitchings R, et al. Adult-onset primary open-angle glaucoma caused by mutations in optineurin. Science. 2002;295(5557):1077-9.

60. Chalasani ML, Radha V, Gupta V, et al. A glaucoma-associated mutant of optineurin selectively induces death of retinal ganglion cells which is inhibited by antioxidants. Investigative ophthalmology & visual science. 2007;48(4):1607-14.

61. Fingert JH, Robin AL, Stone JL, et al. Copy number variations on chromosome 12q14 in patients with normal tension glaucoma. Human molecular genetics. 2011;20(12):2482-94.

62. Awadalla MS, Fingert JH, Roos BE, et al. Copy number variations of TBK1 in Australian patients with primary open-angle glaucoma. American Journal of Ophthalmology. 2015;159(1):124-30.e1.

63. Morton S, Hesson L, Peggie M, et al. Enhanced binding of TBK1 by an optineurin mutant that causes a familial form of primary open angle glaucoma. FEBS letters. 2008;582(6):997-1002.

64. Weil R, Laplantine E, Genin P. Regulation of TBK1 activity by Optineurin contributes to cell cycle-dependent expression of the interferon pathway. Cytokine & growth factor reviews. 2016;29:23-33.

65. Vasiliou V, Gonzalez FJ. Role of CYP1B1 in glaucoma. Annu Rev Pharmacol Toxicol. 2008;48:333-58.

66. van Koolwijk LM, Ramdas WD, Ikram MK, et al. Common genetic determinants of intraocular pressure and primary open-angle glaucoma. PLoS Genet. 2012;8(5):e1002611.

67. Vincent AL, Billingsley G, Buys Y, et al. Digenic inheritance of early-onset glaucoma: CYP1B1, a potential modifier gene. The American Journal of Human Genetics. 2002;70(2):448-60.

68. Gemenetzi M, Yang Y, Lotery AJ. Current concepts on primary open-angle glaucoma genetics: a contribution to disease pathophysiology and future treatment. Eye (Lond). 2012;26(3):355-69.

69. Ramdas WD, van Koolwijk LM, Lemij HG, et al. Common genetic variants associated with open-angle glaucoma. Human molecular genetics. 2011;20(12):2464-71.

70. Sarfarazi M, Child A, Stoilova D, et al. Localization of the fourth locus (GLC1E) for adult-onset primary open-angle glaucoma to the 10p15-p14 region. The American Journal of Human Genetics. 1998;62(3):641-52.

71. Wirtz MK, Samples JR, Rust K, et al. GLC1F, a new primary open-angle glaucoma locus, maps to 7q35-q36. Arch Ophthalmol. 1999;117(2):237-41.

72. Pasutto F, Keller KE, Weisschuh N, et al. Variants in ASB10 are associated with open-angle glaucoma. Human molecular genetics. 2012;21(6):1336-49.

73. Monemi S, Spaeth G, DaSilva A, et al. Identification of a novel adult-onset primary open-angle glaucoma (POAG) gene on 5q22.1. Human molecular genetics. 2005;14(6):725-33.

74. Pasutto F, Matsumoto T, Mardin CY, et al. Heterozygous NTF4 Mutations Impairing Neurotrophin-4 Signaling in Patients with Primary Open-Angle Glaucoma. The American Journal of Human Genetics. 2009;85(4):447-56.

75. Fingert JH, Darbro BW, Qian Q, et al. TBK1 and flanking genes in human retina. Ophthalmic Genet. 2014;35(1):35-40.

76. Nakano M, Ikeda Y, Taniguchi T, et al. Three susceptible loci associated with primary open-angle glaucoma identified by genome-wide association study in a Japanese population.

Proceedings of the National Academy of Sciences of the United States of America. 2009;106(31):12838-42.

77. Thorleifsson G, Walters GB, Hewitt AW, et al. Common variants near CAV1 and CAV2 are associated with primary open-angle glaucoma. Nature genetics. 2010;42(10):906-9.

78. Choquet H, Wiggs JL, Khawaja AP. Clinical implications of recent advances in primary open-angle glaucoma genetics. Eye. 2019.

79. Brodie A, Azaria JR, Ofran Y. How far from the SNP may the causative genes be? Nucleic Acids Research. 2016;44(13):6046-54.

80. Souzeau E, Glading J, Keane M, et al. Predictive genetic testing experience for myocilin primary open-angle glaucoma using the Australian and New Zealand Registry of Advanced Glaucoma. Genetics in medicine : official journal of the American College of Medical Genetics. 2014;16(7):558-63.

81. Wray NR, Goddard ME, Visscher PM. Prediction of individual genetic risk to disease from genome-wide association studies. Genome research. 2007;17(10):1520-8.

82. Khera AV, Chaffin M, Aragam KG, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nature Genetics. 2018;50(9):1219-24.

83. Craig JE, Han X, Qassim A, et al. Multitrait analysis of glaucoma identifies new risk loci and enables polygenic prediction of disease susceptibility and progression. Nat Genet. 2020.

84. Gao XR, Huang H, Kim H. Polygenic Risk Score Is Associated With Intraocular Pressure and Improves Glaucoma Prediction in the UK Biobank CohortGao et al.PRS of IOP in the UK Biobank. Translational Vision Science & Technology. 2019;8(2):10-.

85. Qassim A, Souzeau E, Siggs OM, et al. An Intraocular Pressure Polygenic Risk Score Stratifies Multiple Primary Open-Angle Glaucoma Parameters Including Treatment Intensity. Ophthalmology. 2020.

86. Hysi PG, Cheng CY, Springelkamp H, et al. Genome-wide analysis of multi-ancestry cohorts identifies new loci influencing intraocular pressure and susceptibility to glaucoma. Nature Genetics. 2014;46(10):1126-30.

87. Wiggs JL, Kang JH, Yaspan BL, et al. Common variants near CAV1 and CAV2 are associated with primary open-angle glaucoma in Caucasians from the USA. Human molecular genetics. 2011:ddr382.

88. Loomis SJ, Kang JH, Weinreb RN, et al. Association of CAV1/CAV2 genomic variants with primary open-angle glaucoma overall and by gender and pattern of visual field loss. Ophthalmology. 2014;121(2):508-16.

89. Gharahkhani P, Burdon KP, Fogarty R, et al. Common variants near ABCA1, AFAP1 and GMDS confer risk of primary open-angle glaucoma. Nat Genet. 2014;46(10):1120-5.

90. Osman W, Low S-K, Takahashi A, et al. A genome-wide association study in the Japanese population confirms 9p21 and 14q23 as susceptibility loci for primary open angle glaucoma. Human molecular genetics. 2012:dds103.

91. Burdon KP, Macgregor S, Hewitt AW, et al. Genome-wide association study identifies susceptibility loci for open angle glaucoma at TMCO1 and CDKN2B-AS1. Nat Genet. 2011;43(6):574-8.

92. Burdon KP, Crawford A, Casson RJ, et al. Glaucoma risk alleles at CDKN2B-AS1 are associated with lower intraocular pressure, normal-tension glaucoma, and advanced glaucoma. Ophthalmology. 2012;119(8):1539-45.

93. Gibson J, Griffiths H, De Salvo G, et al. Genome-wide association study of primary open angle glaucoma risk and quantitative traits. Molecular vision. 2012;18:1083-92.

94. Wiggs JL, Yaspan BL, Hauser MA, et al. Common variants at 9p21 and 8q22 are associated with increased susceptibility to optic nerve degeneration in glaucoma. PLoS Genet. 2012;8(4):e1002654.

95. Burdon KP, Mitchell P, Lee A, et al. Association of open-angle glaucoma loci with incident glaucoma in the Blue Mountains Eye Study. Am J Ophthalmol. 2015;159(1):31-6 e1.

96. Cao D, Jiao X, Liu X, et al. CDKN2B polymorphism is associated with primary open-angle glaucoma (POAG) in the Afro-Caribbean population of Barbados, West Indies. PloS one. 2012;7(6):e39278.

97. Fan BJ, Wang DY, Pasquale LR, et al. Genetic variants associated with optic nerve vertical cup-to-disc ratio are risk factors for primary open angle glaucoma in a US Caucasian population. Invest Ophthalmol Vis Sci. 2011;52(3):1788-92.

98. Mabuchi F, Sakurada Y, Kashiwagi K, et al. Association between genetic variants associated with vertical cup-to-disc ratio and phenotypic features of primary open-angle glaucoma. Ophthalmology. 2012;119(9):1819-25.

99. Ramdas WD, van Koolwijk LM, Ikram MK, et al. A genome-wide association study of optic disc parameters. PLoS Genet. 2010;6(6):e1000978.

100. Li Z, Allingham RR, Nakano M, et al. A common variant near TGFBR3 is associated with primary open angle glaucoma. Human molecular genetics. 2015;24(13):3880-92.

101. Nakano M, Ikeda Y, Tokuda Y, et al. Common variants in CDKN2B-AS1 associated with optic-nerve vulnerability of glaucoma identified by genome-wide association studies in Japanese. PloS one. 2012;7(3):e33389.

102. Takamoto M, Kaburaki T, Mabuchi A, et al. Common variants on chromosome 9p21 are associated with normal tension glaucoma. PloS one. 2012;7(7):e40107.

103. Springelkamp H, Hohn R, Mishra A, et al. Meta-analysis of genome-wide association studies identifies novel loci that influence cupping and the glaucomatous process. Nat Commun. 2014;5:4883.

104. Axenovich T, Zorkoltseva I, Belonogova N, et al. Linkage and association analyses of glaucoma related traits in a large pedigree from a Dutch genetically isolated population. Journal of medical genetics. 2011:jmedgenet-2011-100436.

105. Philomenadin FS, Asokan R, Viswanathan N, et al. Genetic association of SNPs near ATOH7, CARD10, CDKN2B, CDC7 and SIX1/SIX6 with the endophenotypes of primary open angle glaucoma in Indian Population. PloS one. 2015;10(3).

106. Springelkamp H, Iglesias AI, Cuellar-Partida G, et al. ARHGEF12 influences the risk of glaucoma by increasing intraocular pressure. Human molecular genetics. 2015:ddv027.

107. Iglesias AI, Springelkamp H, van der Linde H, et al. Exome sequencing and functional analyses suggest that SIX6 is a gene involved in an altered proliferation-differentiation balance early in life and optic nerve degeneration at old age. Human molecular genetics. 2014;23(5):1320-32.

108. Carnes MU, Liu YP, Allingham RR, et al. Discovery and functional annotation of SIX6 variants in primary open-angle glaucoma. PLoS Genet. 2014;10(5):e1004372.

109. Sharma S, Burdon KP, Chidlow G, et al. Association of Genetic Variants in the TMCO1 Gene with Clinical Parameters Related to Glaucoma and Characterization of the Protein in the EyeThe TMCO1 Gene and Glaucoma. Investigative ophthalmology & visual science. 2012;53(8):4917-25.

110. Ozel AB, Moroi SE, Reed DM, et al. Genome-wide association study and meta-analysis of intraocular pressure. Human Genetics. 2014;133(1):41-57.

111. Choquet H, Paylakhi S, Kneeland SC, et al. A multiethnic genome-wide association study of primary open-angle glaucoma identifies novel risk loci. Nat Commun. 2018;9(1):2278.

112. Youngblood H, Hauser MA, Liu Y. Update on the genetics of primary open-angle glaucoma. Exp Eye Res. 2019;188:107795.

113. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. Nature. 2009;461(7265):747-53.

114. Gloss BS, Dinger ME. Realizing the significance of noncoding functionality in clinical genomics. Experimental & Molecular Medicine. 2018;50(8):97.

115. Schierding W, Cutfield WS, O'Sullivan JM. The missing story behind Genome Wide Association Studies: single nucleotide polymorphisms in gene deserts have a story to tell. Frontiers in genetics. 2014;5:39-.

116. Maher B. Personal genomes: The case of the missing heritability. Nature. 2008;456(7218):18-21.

117. Blanco-Gómez A, Castillo-Lluva S, del Mar Sáez-Freire M, et al. Missing heritability of complex diseases: Enlightenment by genetic variants from intermediate phenotypes. BioEssays. 2016;38(7):664-73.

118. Brunham LR, Hayden MR. Hunting human disease genes: lessons from the past, challenges for the future. Hum Genet. 2013;132(6):603-17.

119. Gibson G. Rare and common variants: twenty arguments. Nature Reviews Genetics. 2012;13(2):135-45.

120. Wainschtein P, Jain DP, Yengo L, et al. Recovery of trait heritability from whole genome sequence data. bioRxiv. 2019:588020.

121. Tabor HK, Risch NJ, Myers RM. Candidate-gene approaches for studying complex genetic traits: practical considerations. Nature Reviews Genetics. 2002;3(5):391-7.

122. Mitchell KJ. What is complex about complex disorders? Genome biology. 2012;13(1):237.

123. Bomba L, Walter K, Soranzo N. The impact of rare and low-frequency genetic variants in common disease. Genome biology. 2017;18(1):77.

124. Saint Pierre A, Genin E. How important are rare variants in common disease? Brief Funct Genomics. 2014;13(5):353-61.

125. Zhu H, Wang Z, Wang X, et al. A novel statistical method for rare-variant association studies in general pedigrees. BMC Proc. 2016;10(Suppl 7):193-6.

126. van Koolwijk LM, Bunce C, Viswanathan AC. Gene finding in primary open-angle glaucoma. J Glaucoma. 2013;22(6):473-86.

127. Gottesman, II, Gould TD. The endophenotype concept in psychiatry: etymology and strategic intentions. Am J Psychiatry. 2003;160(4):636-45.

128. Asefa NG, Neustaeter A, Jansonius NM, et al. Heritability of glaucoma and glaucoma-related endophenotypes: Systematic review and meta-analysis. Survey of ophthalmology. 2019;64(6):835-51.

129. Nag A, Venturini C, Small KS, et al. A genome-wide association study of intra-ocular pressure suggests a novel association in the gene FAM125b in the TwinsUK cohort. Human molecular genetics. 2014;23(12):3343-8.

130. Venturini C, Nag A, Hysi PG, et al. Clarifying the role of ATOH7 in glaucoma endophenotypes. British Journal of Ophthalmology. 2014;98(4):562-6.

131. Gharahkhani P, Burdon KP, Cooke Bailey JN, et al. Analysis combining correlated glaucoma traits identifies five new risk loci for open-angle glaucoma. Scientific reports. 2018;8(1):3124.

132. Nannini DR, Torres M, Chen YI, et al. A Genome-Wide Association Study of Vertical Cup-Disc Ratio in a Latino Population. Invest Ophthalmol Vis Sci. 2017;58(1):87-95.

133. Matovinovic E, Kho PF, Lea RA, et al. Genome-wide linkage and association analysis of primary open-angle glaucoma endophenotypes in the Norfolk Island isolate. Molecular vision. 2017;23:660-5.

134. Springelkamp H, Mishra A, Hysi PG, et al. Meta-analysis of Genome-Wide Association Studies Identifies Novel Loci Associated With Optic Disc Morphology. Genetic Epidemiology. 2015;39(3):207-16.

135. Khawaja AP, Cooke Bailey JN, Wareham NJ, et al. Genome-wide analyses identify 68 new loci associated with intraocular pressure and improve risk prediction for primary open-angle glaucoma. Nat Genet. 2018;50(6):778-82.

136. Springelkamp H, Iglesias AI, Mishra A, et al. New insights into the genetics of primary open-angle glaucoma based on meta-analyses of intraocular pressure and optic disc characteristics. Human molecular genetics. 2017;26(2):438-53.

137. Goodwin S, McPherson JD, McCombie WR. Coming of age: ten years of next-generation sequencing technologies. Nat Rev Genet. 2016;17(6):333-51.

138. Rabbani B, Tekin M, Mahdieh N. The promise of whole-exome sequencing in medical genetics. J Hum Genet. 2014;59(1):5-15.

139. Majewski J, Schwartzentruber J, Lalonde E, et al. What can exome sequencing do for you? J Med Genet. 2011;48(9):580-9.

140. Cirulli ET, Goldstein DB. Uncovering the roles of rare variants in common disease through whole-genome sequencing. Nature Reviews Genetics. 2010;11(6):415-25.

141. Altmann A, Weber P, Bader D, et al. A beginners guide to SNP calling from high-throughput DNA-sequencing data. Human Genetics. 2012;131(10):1541-54.

142. Dolled-Filhart MP, Lee M, Ou-yang C-w, et al. Computational and bioinformatics frameworks for next-generation whole exome and genome sequencing. The Scientific World Journal. 2013;2013.

143. Chen F, Klein AP, Klein BEK, et al. Exome array analysis identifies CAV1/CAV2 as a susceptibility locus for intraocular pressure. Investigative Ophthalmology and Visual Science. 2015;56(1):544-51.

144. Liu T, Xie L, Ye J, et al. Family-based analysis identified CD2 as a susceptibility gene for primary open angle glaucoma in Chinese Han population. Journal of cellular and molecular medicine. 2014;18(4):600-9.

145. Mackay DS, Bennett TM, Shiels A. Exome sequencing identifies a missense variant in efemp1 co-segregating in a family with autosomal dominant primary open-angle glaucoma. PloS one. 2015;10(7).

146. Micheal S, Saksens NT, Hogewind BF, et al. Identification of TP53BP2 as a Novel Candidate Gene for Primary Open Angle Glaucoma by Whole Exome Sequencing in a Large Multiplex Family. Mol Neurobiol. 2017.

147. Zhou T, Souzeau E, Sharma S, et al. Rare variants in optic disc area gene CARD10 enriched in primary open-angle glaucoma. Molecular genetics & genomic medicine. 2016;4(6):624-33.

148. Micheal S, Hogewind BF, Khan MI, et al. Variants in the PRPF8 Gene are Associated with Glaucoma. Mol Neurobiol. 2017.

149. Zhou T, Souzeau E, Sharma S, et al. Whole exome sequencing implicates eye development, the unfolded protein response and plasma membrane homeostasis in primary open-angle glaucoma. PloS one. 2017;12(3):e0172427.

150. Bahlo M, Bromhead CJ. Generating linkage mapping files from Affymetrix SNP chip data. Bioinformatics. 2009;25(15):1961-2.

151. Bahlo M, Tankard R, Lukic V, et al. Using familial information for variant filtering in high-throughput sequencing studies. Hum Genet. 2014;133(11):1331-41.

152. Han L, Abney M. Identity by descent estimation with dense genome-wide genotype data. Genet Epidemiol. 2011;35(6):557-67.

153. Wirtz MK, Samples JR, Kramer PL, et al. Mapping a gene for adult-onset primary open-angle glaucoma to chromosome 3q. American journal of human genetics. 1997;60(2):296-304.

154. Mackey D, Craig J, editors. Glaucoma Inheritance Study in Tasmania: an international collaboration. American Academy of Ophthalmology Basic Sciences Course Section 13 International Ophthalmology, Part 5 Collaborative Research; 2002.

155. Coote MA, McCartney PJ, Wilkinson RM, et al. The 'GIST' score: ranking glaucoma for genetic studies. Glaucoma Inheritance Study of Tasmania. Ophthalmic Genet. 1996;17(4):199-208.

156. Kelly BJ, Fitch JR, Hu Y, et al. Churchill: an ultra-fast, deterministic, highly scalable and balanced parallelization strategy for the discovery of human genetic variation in clinical and population-scale genomics. Genome biology. 2015;16(1):6.

157. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. 2013.

158. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome research. 2010;20(9):1297-303.

159. Poplin R, Ruano-Rubio V, DePristo MA, et al. Scaling accurate genetic variant discovery to tens of thousands of samples. bioRxiv. 2018:201178.

160. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010;38(16):e164.

161. Yang H, Wang K. Genomic variant annotation and prioritization with ANNOVAR and wANNOVAR. Nat Protoc. 2015;10(10):1556-66.

162. Almasy L, Blangero J. Multipoint quantitative-trait linkage analysis in general pedigrees. American journal of human genetics. 1998;62(5):1198-211.

163. Boerwinkle E, Chakraborty R, Sing C. The use of measured genotype information in the analysis of quantitative phenotypes in man. Annals of human genetics. 1986;50(2):181-94.

164. Boerwinkle E, Sing C. The use of measured genotype information in the analysis of quantitative phenotypes in man. Annals of human genetics. 1987;51(3):211-26.

165. Saint Pierre A, Vitezica Z, Martinez M. A comparative study of three methods for detecting association of quantitative traits in samples of related subjects. BMC Proceedings. 2009;3(Suppl 7):S122-S.

166. Almasy L, Warren DM. Software for quantitative trait analysis. Human genomics. 2005;2(3):1.

167. Almasy L, Blangero J. Variance Component Methods for Analysis of Complex Phenotypes. Cold Spring Harbor protocols. 2010;2010(5):pdb.top77-pdb.top.

168. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. 2013.

169. Wickham H. ggplot2. WIREs Comput Stat. 2011;3(2):180-5.

170. Mäkinen V-P, Parkkonen M, Wessman M, et al. High-throughput pedigree drawing. European Journal of Human Genetics. 2005;13(8):987-9.

171. Choquet H, Thai KK, Yin J, et al. A large multi-ethnic genome-wide association study identifies novel genetic loci for intraocular pressure. Nat Commun. 2017;8(1):2108.

172. Hoffmann TJ, Tang H, Thornton TA, et al. Genome-wide association and admixture analysis of glaucoma in the Women's Health Initiative. Human molecular genetics. 2014;23(24):6634-43.

173. Trifan OC, Traboulsi EI, Stoilova D, et al. A third locus (GLC1D) for adult-onset primary open-angle glaucoma maps to the 8q23 region. American journal of ophthalmology. 1998;126(1):17-28.

174. Baird PN, Foote SJ, Mackey DA, et al. Evidence for a novel glaucoma locus at chromosome 3p21-22. Hum Genet. 2005;117(2-3):249-57.

175. Stoilova D, Child A, Trifan OC, et al. Localization of a locus (GLC1B) for adult-onset primary open angle glaucoma to the 2cen–q13 region. Genomics. 1996;36(1):142-50.

176. Suriyapperuma SP, Child A, Desai T, et al. A new locus (GLC1H) for adult-onset primary open-angle glaucoma maps to the 2p15-p16 region. Archives of Ophthalmology. 2007;125(1):86-92.

177. Allingham RR, Wiggs JL, Hauser ER, et al. Early adult-onset POAG linked to 15q11-13 using ordered subset analysis. Investigative ophthalmology & visual science. 2005;46(6):2002-5.

178. Coleman AL, Miglior S. Risk factors for glaucoma onset and progression. Survey of ophthalmology. 2008;53 Suppl1:S3-10.

179. Iglesias AI, Mishra A, Vitart V, et al. Cross-ancestry genome-wide association analysis of corneal thickness strengthens link between complex and Mendelian eye diseases. Nat Commun. 2018;9(1):1864.

180. Garway-Heath DF, Ruben ST, Viswanathan A, et al. Vertical cup/disc ratio in relation to optic disc size: its value in the assessment of the glaucoma suspect. Br J Ophthalmol. 1998;82(10):1118-24.

181. Tunny TJ, Richardson KA, Clark CV, et al. The atrial natriuretic peptide gene in patients with familial primary open-angle glaucoma. Biochemical and biophysical research communications. 1996;223(2):221-5.

182. Micheal S, Ayub H, Zafar SN, et al. Identification of novel CYP1B1 gene mutations in patients with primary congenital and primary open-angle glaucoma. Clinical and Experimental Ophthalmology. 2015;43(1):31-9.

183. Wang Z, Li M, Li L, et al. Association of single nucleotide polymorphisms in the CYP1B1 gene with the risk of primary open-angle glaucoma: A meta-analysis. Genetics and Molecular Research. 2015;14(4):17262-72.

184. Wang CY, Shen YC, Lo FY, et al. Polymorphism in the IL-1α (-889) locus associated with elevated risk of primary open angle glaucoma. Molecular vision. 2006;12:1380-5.

185. Karczewski KJ, Francioli LC, Tiao G, et al. Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. bioRxiv. 2019:531210.

186. Inagaki Y, Mashima Y, Fuse N, et al. Polymorphism of β-adrenergic receptors and susceptibility of open-angle glaucoma. Molecular vision. 2006;12:673-80.

187. Nemesure B, Jiao X, He Q, et al. A genome-wide scan for primary open-angle glaucoma (POAG): the Barbados Family Study of Open-Angle Glaucoma. Human genetics. 2003;112(5-6):600-9.

188. Duggal P, Klein AP, Lee KE, et al. A genetic contribution to intraocular pressure: the beaver dam eye study. Invest Ophthalmol Vis Sci. 2005;46(2):555-60.

189. Charlesworth JC, Dyer TD, Stankovich JM, et al. Linkage to 10q22 for maximum intraocular pressure and 1p32 for maximum cup-to-disc ratio in an extended primary open-angle glaucoma pedigree. Invest Ophthalmol Vis Sci. 2005;46(10):3723-9.

190. Wiggs JL, Allingham RR, Hossain A, et al. Genome-wide scan for adult onset primary open angle glaucoma. Human molecular genetics. 2000;9(7):1109-17.

191. Duggal P, Klein AP, Lee KE, et al. Identification of novel genetic loci for intraocular pressure: a genomewide scan of the Beaver Dam Eye Study. Arch Ophthalmol. 2007;125(1):74-9.

192. Doughty MJ, Zaman ML. Human corneal thickness and its impact on intraocular pressure measures: a review and meta-analysis approach. Survey of ophthalmology. 2000;44(5):367-408.

193. Pasutto F, Chavarria-Soley G, Mardin CY, et al. Heterozygous Loss-of-Function Variants in CYP1B1 Predispose to Primary Open-Angle Glaucoma. Investigative Ophthalmology & Visual Science. 2010;51(1):249-54.

194. Abu-Amero KK, Morales J, Aljasim LA, et al. CYP1B1 Mutations are a Major Contributor to Juvenile-Onset Open Angle Glaucoma in Saudi Arabia. Ophthalmic Genet. 2015;36(2):184-7.

195. Bayat B, Yazdani S, Alavi A, et al. Contributions of MYOC and CYP1B1 mutations to JOAG. Molecular vision. 2008;14:508-17.

196. Ulmer M, Li J, Yaspan BL, et al. Genome-Wide Analysis of Central Corneal Thickness in Primary Open-Angle Glaucoma Cases in the NEIGHBOR and GLAUGEN ConsortiaThe Effects of CCT-Associated Variants on POAG Risk. Investigative Ophthalmology & Visual Science. 2012;53(8):4468-74.

197. Gao X, Nannini DR, Corrao K, et al. Genome-wide association study identifies WNT7B as a novel locus for central corneal thickness in Latinos. Human molecular genetics. 2016.

198. Hoehn R, Zeller T, Verhoeven VJ, et al. Population-based meta-analysis in Caucasians confirms association with COL5A1 and ZNF469 but not COL8A2 with central corneal thickness. Human genetics. 2012;131(11):1783-93.

199. Lu Y, Dimasi DP, Hysi PG, et al. Common genetic variants near the Brittle Cornea Syndrome locus ZNF469 influence the blinding disease risk factor central corneal thickness. PLoS Genet. 2010;6(5):e1000947.

200. Lu Y, Vitart V, Burdon KP, et al. Genome-wide association analyses identify multiple loci associated with central corneal thickness and keratoconus. Nat Genet. 2013;45(2):155-63.

201. Vitart V, Benčić G, Hayward C, et al. New loci associated with central cornea thickness include COL5A1, AKAP13 and AVGR8. Human molecular genetics. 2010;19(21):4304-11.

202. Diskin S, Kumar J, Cao Z, et al. Detection of differentially expressed glycogenes in trabecular meshwork of eyes with primary open-angle glaucoma. Invest Ophthalmol Vis Sci. 2006;47(4):1491-9.

203. Weinreb RN, Leung CK, Crowston JG, et al. Primary open-angle glaucoma. Nature reviews Disease primers. 2016;2:16067.

204. Sud A, Del Bono E, Haines J, et al. Fine mapping of the GLC1K juvenile primary open-angle glaucoma locus and exclusion of candidate genes. 2008.

205. Fan BJ, Ko WC, Wang DY, et al. Fine mapping of new glaucoma locus GLC1M and exclusion of neuregulin 2 as the causative gene. Molecular vision. 2007;13:779-84.

206. Gordon MO, Beiser JA, Brandt JD, et al. The Ocular Hypertension Treatment Study: baseline factors that predict the onset of primary open-angle glaucoma. Arch Ophthalmol. 2002;120(6):714-20; discussion 829-30.

207. Gordon MO, Torri V, Miglior S, et al. Validated prediction model for the development of primary open-angle glaucoma in individuals with ocular hypertension. Ophthalmology. 2007;114(1):10-9.

208. Chen Z, Tan K-R, Bull SB. Multiphase analysis by linkage, quantitative transmission disequilibrium, and measured genotype: systolic blood pressure in complex Mexican American pedigrees. BMC Proceedings. 2014;8(Suppl 1):S108-S.

209. Scheet P, Stephens M. A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. American journal of human genetics. 2006;78(4):629-44.

210. Marchini J, Howie B. Genotype imputation for genome-wide association studies. Nature reviews Genetics. 2010;11(7):499-511.

211. Glahn DC, Nimgaonkar VL, Raventos H, et al. Rediscovering the value of families for psychiatric genetics research. Molecular psychiatry. 2018.

212. Blangero J, Williams JT, Almasy L. Variance component methods for detecting complex trait loci. Adv Genet. 2001;42:151-81.

213. Dyer TD, Blangero J, Williams JT, et al. The effect of pedigree complexity on quantitative trait linkage analysis. Genet Epidemiol. 2001;21 Suppl 1:S236-43.

214. Ekstrom CT. Power of multipoint identity-by-descent methods to detect linkage using variance component models. Genet Epidemiol. 2001;21(4):285-98.

215. Kruglyak L. The use of a genetic map of biallelic markers in linkage studies. Nat Genet. 1997;17(1):21-4.

216. Thalamuthu A, Mukhopadhyay I, Ray A, et al. A comparison between microsatellite and single-nucleotide polymorphism markers with respect to two measures of information content. BMC genetics. 2005;6 Suppl 1(Suppl 1):S27-S.

217. Dunn G, Hinrichs AL, Bertelsen S, et al. Microsatellites versus single-nucleotide polymorphisms in linkage analysis for quantitative and qualitative measures. BMC genetics. 2005;6 Suppl 1(Suppl 1):S122-S.

218. Schaid DJ, Guenther JC, Christensen GB, et al. Comparison of microsatellites versus single-nucleotide polymorphisms in a genome linkage screen for prostate cancer-susceptibility Loci. American journal of human genetics. 2004;75(6):948-65.

219. Gazal S, Gosset S, Verdura E, et al. Can whole-exome sequencing data be used for linkage analysis? Eur J Hum Genet. 2016;24(4):581-6.

220. Ott J, Wang J, Leal SM. Genetic linkage analysis in the age of whole-genome sequencing. Nat Rev Genet. 2015;16(5):275-84.

221. Smith KR, Bromhead CJ, Hildebrand MS, et al. Reducing the exome search space for mendelian diseases using genetic linkage analysis of exome genotypes. Genome biology. 2011;12(9):R85.

222. Browning BL, Browning SR. Improving the accuracy and efficiency of identity-by-descent detection in population data. Genetics. 2013;194(2):459-71.

223. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. American journal of human genetics. 2007;81(3):559-75.

224. O'Connell JR, Weeks DE. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. American journal of human genetics. 1998;63(1):259-66.

225. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078-9.

226. Han L, Abney M. Using identity by descent estimation with dense genotype data to detect positive selection. European Journal of Human Genetics. 2013;21(2):205-11.

227. Jacquard A. Genetic Information Given by a Relative. Biometrics. 1972;28(4):1101-14.

228. Peralta JM, Blackburn NB, Porto A, et al. Genome-wide linkage scan for loci influencing plasma triglyceride levels. BMC Proceedings. 2018;12(9):52.

229. Blackburn NB, Michael LF, Meikle PJ, et al. Rare DEGS1 variant significantly alters de novo ceramide synthesis pathway. Journal of lipid research. 2019;60(9):1630-9.

230. Wensor MD, McCarty CA, Stanislavsky YL, et al. The prevalence of glaucoma in the Melbourne Visual Impairment Project. Ophthalmology. 1998;105(4):733-9.

231. Blangero J, Williams JT, Almasy L. Robust LOD scores for variance component-based linkage analysis. Genet Epidemiol. 2000;19 Suppl 1:S8-14.

232. Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nat Genet. 1995;11(3):241-7.

233. Crooks KR, Allingham RR, Qin X, et al. Genome-wide linkage scan for primary open angle glaucoma: influences of ancestry and age at diagnosis. PloS one. 2011;6(7):e21967.

234. Eggers S, Smith KR, Bahlo M, et al. Whole exome sequencing combined with linkage analysis identifies a novel 3 bp deletion in NR5A1. European Journal of Human Genetics. 2015;23(4):486-93.

235. Hildebrand MS, Tankard R, Gazina EV, et al. PRIMA1 mutation: a new cause of nocturnal frontal lobe epilepsy. Annals of Clinical and Translational Neurology. 2015;2(8):821-30.

236. Guergueltcheva V, Azmanov DN, Angelicheva D, et al. Autosomal-recessive congenital cerebellar ataxia is caused by mutations in metabotropic glutamate receptor 1. American journal of human genetics. 2012;91(3):553-64.

237. Chen F, Szymanski EP, Olivier KN, et al. Whole-Exome Sequencing Identifies the 6q12-q16 Linkage Region and a Candidate Gene, TTK, for Pulmonary Nontuberculous Mycobacterial Disease. American journal of respiratory and critical care medicine. 2017;196(12):1599-604.

238. Norton N, Li D, Rampersaud E, et al. Exome sequencing and genome-wide linkage analysis in 17 families illustrate the complex contribution of TTN truncating variants to dilated cardiomyopathy. Circulation Cardiovascular genetics. 2013;6(2):144-53.

239. Ng SB, Turner EH, Robertson PD, et al. Targeted capture and massively parallel sequencing of 12 human exomes. Nature. 2009;461(7261):272-6.

240. Sakharkar MK, Chow VT, Kangueane P. Distributions of exons and introns in the human genome. In Silico Biol. 2004;4(4):387-93.

241. Blackburn NB, Porto A, Peralta JM, et al. Heritability and genetic associations of triglyceride and HDL-C levels using pedigree-based and empirical kinships. BMC Proc. 2018;12(Suppl 9):34.

242. van der Valk R, Webers CA, Schouten JS, et al. Intraocular pressure-lowering effects of all commonly used glaucoma drugs: a meta-analysis of randomized clinical trials. Ophthalmology. 2005;112(7):1177-85.

243. Ozer MA, Acar M, Yildirim C. Intraocular pressure-lowering effects of commonly used fixed combination drugs with timolol in the management of primary open angle glaucoma. International journal of ophthalmology. 2014;7(5):832-6.

244. Williams JT, Blangero J. Power of variance component linkage analysis to detect quantitative trait loci. Ann Hum Genet. 1999;63(Pt 6):545-63.

245. Brisbin A, Weissman M, Fyer A, et al. Bayesian Linkage Analysis of Categorical Traits for Arbitrary Pedigree Designs. PloS one. 2010;5:e12307.

246. Almasy L, Blangero J. Exploring positional candidate genes: linkage conditional on measured genotype. Behavior genetics. 2004;34(2):173-7.

247. Nyholt DR. All LODs are not created equal. American journal of human genetics. 2000;67(2):282-8.

248. MacGregor S, Ong J-S, An J, et al. Genome-wide association study of intraocular pressure uncovers new pathways to glaucoma. Nature Genetics. 2018;50(8):1067-71.

249. Wagner AH, Anand VN, Wang WH, et al. Exon-level expression profiling of ocular tissues. Exp Eye Res. 2013;111:105-11.

250. Kramer A, Green J, Pollard J, Jr., et al. Causal analysis approaches in Ingenuity Pathway Analysis. Bioinformatics. 2014;30(4):523-30.

251. Danford ID, Verkuil LD, Choi DJ, et al. Characterizing the "POAGome": A bioinformatics-driven approach to primary open-angle glaucoma. Prog Retin Eye Res. 2017;58:89-114.

252. Cubillos-Rojas M, Amair-Pinedo F, Peiro-Jordan R, et al. The E3 ubiquitin protein ligase HERC2 modulates the activity of tumor protein p53 by regulating its oligomerization. The Journal of biological chemistry. 2014;289(21):14782-95.

253. Grikscheit K, Frank T, Wang Y, et al. Junctional actin assembly is mediated by Formin-like 2 downstream of Rac1. The Journal of cell biology. 2015;209(3):367-76.

254. Hatzfeld M, Green KJ, Sauter H. Targeting of p0071 to desmosomes and adherens junctions is mediated by different protein domains. Journal of cell science. 2003;116(Pt 7):1219-33.

255. Bekker-Jensen S, Rendtlew Danielsen J, Fugger K, et al. HERC2 coordinates ubiquitin-dependent assembly of DNA repair factors on damaged chromosomes. Nature cell biology. 2010;12(1):80-6; sup pp 1-12.

256. Eiberg H, Troelsen J, Nielsen M, et al. Blue eye color in humans may be caused by a perfectly associated founder mutation in a regulatory element located within the HERC2 gene inhibiting OCA2 expression. Hum Genet. 2008;123(2):177-87.

257. Sturm RA, Duffy DL, Zhao ZZ, et al. A single SNP in an evolutionary conserved region within intron 86 of the HERC2 gene determines human blue-brown eye color. American journal of human genetics. 2008;82(2):424-31.

258. Sulem P, Gudbjartsson DF, Stacey SN, et al. Genetic determinants of hair, eye and skin pigmentation in Europeans. Nat Genet. 2007;39(12):1443-52.

259. Puffenberger EG, Jinks RN, Wang H, et al. A homozygous missense mutation in HERC2 associated with global developmental delay and autism spectrum disorder. Human mutation. 2012;33(12):1639-46.

260. Harlalka GV, Baple EL, Cross H, et al. Mutation of HERC2 causes developmental delay with Angelman-like features. J Med Genet. 2013;50(2):65-73.

261. Lin HJ, Chen WC, Tsai FJ, et al. Distributions of p53 codon 72 polymorphism in primary open angle glaucoma. British Journal of Ophthalmology. 2002;86(7):767-70.

262. Neamatzadeh H, Soleimanizad R, Zare-Shehneh M, et al. Association between p53 codon 72 (Arg72pro) polymorphism and primary open-angle glaucoma in Iranian patients. Iranian biomedical journal. 2015;19(1):51-6.

263. Daugherty CL, Curtis H, Realini T, et al. Primary open angle glaucoma in a Caucasian population is associated with the p53 codon 72 polymorphism. Molecular vision. 2009;15:1939-44.

264. Fan BJ, Liu K, Wang DY, et al. Association of polymorphisms of tumor necrosis factor and tumor protein p53 with primary open-angle glaucoma. Invest Ophthalmol Vis Sci. 2010;51(8):4110-6.

265. Acharya M, Mitra S, Mukhopadhyay A, et al. Distribution of p53 codon 72 polymorphism in Indian primary open angle glaucoma patients. Molecular vision. 2002;8:367-71.

266. Dimasi DP, Hewitt AW, Green CM, et al. Lack of association of p53 polymorphisms and haplotypes in high and normal tension open angle glaucoma. J Med Genet. 2005;42(9):e55.

267. Olivier M, Hollstein M, Hainaut P. TP53 mutations in human cancers: origins, consequences, and clinical use. Cold Spring Harb Perspect Biol. 2010;2(1):a001008-a.

268. Cubillos-Rojas M, Schneider T, Bartrons R, et al. NEURL4 regulates the transcriptional activity of tumor suppressor protein p53 by modulating its oligomerization. Oncotarget. 2017;8(37):61824-36.

269. Lehtokari VL, Kiiski K, Sandaradura SA, et al. Mutation update: the spectra of nebulin variants and associated myopathies. Human mutation. 2014;35(12):1418-26.

270. Ahram DF, Grozdanic SD, Kecova H, et al. Variants in Nebulin (NEB) Are Linked to the Development of Familial Primary Angle Closure Glaucoma in Basset Hounds. PloS one. 2015;10(5):e0126660.

271. Webber HC, Bermudez JY, Millar JC, et al. The Role of Wnt/β-Catenin Signaling and K-Cadherin in the Regulation of Intraocular Pressure. Investigative ophthalmology & visual science. 2018;59(3):1454-66.

272. Wecker T, Han H, Borner J, et al. Effects of TGF-beta2 on cadherins and beta-catenin in human trabecular meshwork cells. Invest Ophthalmol Vis Sci. 2013;54(10):6456-62.

273. Wang DY, Ray A, Rodgers K, et al. Global Gene Expression Changes in Rat Retinal Ganglion Cells in Experimental Glaucoma. Investigative Ophthalmology & Visual Science. 2010;51(8):4084-95.

274. Keil R, Schulz J, Hatzfeld M. p0071/PKP4, a multifunctional protein coordinating cell adhesion with cytoskeletal organization. Biol Chem. 2013;394(8):1005-17.

275. Nelson WJ, Nusse R. Convergence of Wnt, beta-catenin, and cadherin pathways. Science (New York, NY). 2004;303(5663):1483-7.

276. Tian H, Sanders E, Reynolds A, et al. Ocular anterior segment dysgenesis upon ablation of p120 catenin in neural crest cells. Invest Ophthalmol Vis Sci. 2012;53(9):5139-53.

277. Kramer PR, Wray S. Novel gene expressed in nasal region influences outgrowth of olfactory axons and migration of luteinizing hormone-releasing hormone (LHRH) neurons. Genes Dev. 2000;14(14):1824-34.

278. Hradsky J, Bernstein HG, Marunde M, et al. Alternative splicing, expression and cellular localization of Calneuron-1 in the rat and human brain. The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society. 2015;63(10):793-804.

279. Sewduth RN, Jaspard-Vinassa B, Peghaire C, et al. The ubiquitin ligase PDZRN3 is required for vascular morphogenesis through Wnt/planar cell polarity signalling. Nat Commun. 2014;5:4832.

280. Honda T, Inui M. PDZRN3 regulates differentiation of myoblasts into myotubes through transcriptional and posttranslational control of Id2. J Cell Physiol. 2019;234(3):2963-72.

281. Honda T, Ishii A, Inui M. Regulation of adipocyte differentiation of 3T3-L1 cells by PDZRN3. American journal of physiology Cell physiology. 2013;304(11):C1091-7.

282. Honda T, Yamamoto H, Ishii A, et al. PDZRN3 negatively regulates BMP-2-induced osteoblast differentiation through inhibition of Wnt signaling. Mol Biol Cell. 2010;21(18):3269-77.

283. Sugrue LP, Desikan RS. What Are Polygenic Scores and Why Are They Important? JAMA. 2019;321(18):1820-1.

284. de Jong S, Diniz MJA, Saloma A, et al. Applying polygenic risk scoring for psychiatric disorders to a large family with bipolar disorder and major depressive disorder. Communications biology. 2018;1(1):163.

285. Browning SR, Browning BL. Identity by descent between distant relatives: detection and applications. Annual review of genetics. 2012;46:617-33.

286. Saint-Pierre A, D'Elia Y, Ciullo M, et al. SNP-based linkage analysis in extended pedigrees: comparison between two alternative approaches. Hum Hered. 2014;78(1):27-37.

287. Blangero J, Williams JT, Almasy L. Novel family-based approaches to genetic risk in thrombosis. Journal of thrombosis and haemostasis : JTH. 2003;1(7):1391-7.

288. Liu Y, Bailey JC, Helwa I, et al. A Common Variant in MIR182 Is Associated With Primary Open-Angle Glaucoma in the NEIGHBORHOOD Consortium. Invest Ophthalmol Vis Sci. 2016;57(10):3974-81.

289. Ghanbari M, Iglesias AI, Springelkamp H, et al. A Genome-Wide Scan for MicroRNA-Related Genetic Variants Associated With Primary Open-Angle Glaucoma. Invest Ophthalmol Vis Sci. 2017;58(12):5368-77.

290. Zhang F, Lupski JR. Non-coding genetic variants in human disease. Human molecular genetics. 2015;24(R1):R102-10.

291. Sanger F, Coulson AR, Barrell BG, et al. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. Journal of molecular biology. 1980;143(2):161-78.

292. Souzeau E, Goldberg I, Healey PR, et al. Australian and New Zealand Registry of Advanced Glaucoma: methodology and recruitment. Clinical & experimental ophthalmology. 2012;40(6):569-75.

293. Sayyad Z, Sirohi K, Radha V, et al. 661W is a retinal ganglion precursor-like cell line in which glaucoma-associated optineurin mutants induce cell death selectively. Scientific reports. 2017;7(1):16855-.

294. Swarup G, Sayyad Z. Altered Functions and Interactions of Glaucoma-Associated Mutants of Optineurin. Frontiers in Immunology. 2018;9(1287).

295. Zhao L, Liu H, Yuan X, et al. Comparative study of whole exome sequencing-based copy number variation detection tools. BMC Bioinformatics. 2020;21(1):97.

296. Auer PL, Johnsen JM, Johnson AD, et al. Imputation of exome sequence variants into population- based samples and blood-cell-trait-associated loci in African Americans: NHLBI GO Exome Sequencing Project. American journal of human genetics. 2012;91(5):794-808.

297. Zhou T, Souzeau E, Siggs OM, et al. Contribution of Mutations in Known Mendelian Glaucoma Genes to Advanced Early-Onset Primary Open-Angle Glaucoma. Invest Ophthalmol Vis Sci. 2017;58(3):1537-44.

298. Souzeau E, Hayes M, Zhou T, et al. Occurrence of CYP1B1 Mutations in Juvenile Open-Angle Glaucoma With Advanced Visual Field Loss. JAMA Ophthalmol. 2015;133(7):826-33.

299. Song N, Leng L, Yang X-J, et al. Compound heterozygous mutations in CYP1B1 gene leads to severe primary congenital glaucoma phenotype. International journal of ophthalmology. 2019;12(6):909-14.

300. Zhan X, Hu Y, Li B, et al. RVTESTS: an efficient and comprehensive tool for rare variant association analysis using sequence data. Bioinformatics. 2016;32(9):1423-6.

301. Chen H, Meigs JB, Dupuis J. Sequence kernel association test for quantitative traits in family samples. Genetic epidemiology. 2013;37(2):196-204.

302. Wang X, Zhang Z, Morris N, et al. Rare variant association test in family-based sequencing studies. Briefings in Bioinformatics. 2016;18(6):954-61.

303. Jiang D, McPeek MS. Robust rare variant association testing for quantitative traits in samples with related individuals. Genetic epidemiology. 2014;38(1):10-20.

304. Gharahkhani P, Jorgenson E, Hysi P, et al. A large cross-ancestry meta-analysis of genome-wide association studies identifies 69 novel risk loci for primary open-angle glaucoma and includes a genetic link with Alzheimer's disease. bioRxiv. 2020:2020.01.30.927822.

305. Khawaja AP, Viswanathan AC. Are we ready for genetic testing for primary open-angle glaucoma? Eye (Lond). 2018;32(5):877-83.

306. Ran FA, Hsu PD, Wright J, et al. Genome engineering using the CRISPR-Cas9 system. Nat Protoc. 2013;8(11):2281-308.

307. Jain A, Zode G, Kasetti RB, et al. CRISPR-Cas9-based treatment of myocilin-associated glaucoma. Proceedings of the National Academy of Sciences of the United States of America. 2017;114(42):11199-204.

308. Wu J, Bell OH, Copland DA, et al. Gene Therapy for Glaucoma by Ciliary Body Aquaporin 1 Disruption Using CRISPR-Cas9. Molecular Therapy. 2020;28(3):820-9.

309. Liu W, Liu Y, Qin XJ, et al. AQP1 and SLC4A10 as candidate genes for primary open-angle glaucoma. Molecular vision. 2010;16:93-7.

310. Zhao Y, Zhu H, Yang Y, et al. AQP1 suppression by ATF4 triggers trabecular meshwork tissue remodelling in ET-1-induced POAG. Journal of cellular and molecular medicine. 2020.

311. Jünemann AG, von Ahsen N, Reulbach U, et al. C677T variant in the methylentetrahydrofolate reductase gene is a genetic risk factor for primary open-angle glaucoma. American journal of ophthalmology. 2005;139(4):721-3.

312. Nilforoushan N, Aghapour S, Raoofian R, et al. Lack of association between the C677T single nucleotide polymorphism of the MTHFR gene and glaucoma in Iranian patients. Acta medica Iranica. 2012;50(3):208-12.

313. Wolf C, Gramer E, Müller-Myhsok B, et al. Evaluation of nine candidate genes in patients with normal tension glaucoma: A case control study. BMC medical genetics. 2009;10:91.

314. Desronvil T, Logan-Wyatt D, Abdrabou W, et al. Distribution of COL8A2 and COL8A1 gene variants in Caucasian primary open angle glaucoma patients with thin central corneal thickness. Molecular vision. 2010;16:2185-91.

315. Vithana EN, Aung T, Khor CC, et al. Collagen-related genes influence the glaucoma risk factor, central corneal thickness. Human molecular genetics. 2011;20(4):649-58.

316. Cornes BK, Khor CC, Nongpiur ME, et al. Identification of four novel variants that influence central corneal thickness in multi-ethnic Asian populations. Human molecular genetics. 2012;21(2):437-45.

317. Khor CC, Ramdas WD, Vithana EN, et al. Genome-wide association studies in Asians confirm the involvement of ATOH7 and TGFBR3, and further identify CARD10 as a novel locus influencing optic disc area. Human molecular genetics. 2011;20(9):1864-72.

318. Juronen E, Tasa G, Veromann S, et al. Polymorphic glutathione S-transferase M1 is a risk factor of primary open-angle glaucoma among Estonians. Experimental Eye Research. 2000;71(5):447-52.

319. Izzotti A, Saccà SC, Cartiglia C, et al. Oxidative deoxyribonucleic acid damage in the eyes of glaucoma patients. The American journal of medicine. 2003;114(8):638-46.

320. Jansson M, Rada A, Tomic L, et al. Analysis of the Glutathione S-transferase M1 gene using pyrosequencing and multiplex PCR-no evidence of association to glaucoma. Experimental Eye Research. 2003;77(2):239-43.

321. Yildirim Ö, Ateş NA, Tamer L, et al. May glutathione S-transferase M1 positive genotype afford protection against primary open-angle glaucoma? Graefe's Archive for Clinical and Experimental Ophthalmology. 2005;243(4):327-33.

322. Ünal M, Güven M, Devranoğlu K, et al. Glutathione S transferase M1 and T1 genetic polymorphisms are related to the risk of primary open-angle glaucoma: A study in a Turkish population. British Journal of Ophthalmology. 2007;91(4):527-30.

323. Abu-Amero KK, Morales J, Mohamed GH, et al. Glutathione S-transferase M1 and T1 polymorphisms in Arab glaucoma patients. 2008.

324. Rocha AV, Talbot T, Magalhães da Silva T, et al. Is the GSTM1 null polymorphism a risk factor in primary open angle glaucoma? 2011.

325. Acharya M, Mookherjee S, Bhattacharjee A, et al. Evaluation of the OPTC gene in primary open angle glaucoma: Functional significance of a silent change. BMC molecular biology. 2007;8.

326. Chen LJ, Tam POS, Leung DYL, et al. SNP rs1533428 at 2p16.3 as a marker for late-onset primary openangle glaucoma. Molecular vision. 2012;18:1629-39.

327. Kim K, Heo DW, Kim S, et al. Expansive marker analysis replicating the association of glaucoma susceptibility with human chromosome loci 1q43 and 10p12.31. European Journal of Human Genetics. 2014;22(3):409-13.

328. Meguro A, Inoko H, Ota M, et al. Genome-wide association study of normal tension glaucoma: common variants in SRBD1 and ELOVL5 contribute to disease susceptibility. Ophthalmology. 2010;117(7):1331-8. e5.

329. Mabuchi F, Sakurada Y, Kashiwagi K, et al. Association between SRBD1 and ELOVL5 gene polymorphisms and primary open-angle glaucoma. Investigative Ophthalmology and Visual Science. 2011;52(7):4626-9.

330. Shi D, Funayama T, Mashima Y, et al. Association of HK2 and NCK2 with Normal Tension Glaucoma in the Japanese Population. PloS one. 2013;8(1).

331. Akiyama M, Yatsu K, Ota M, et al. Microsatellite analysis of the GLC1B locus on chromosome 2 points to NCK2 as a new candidate gene for normal tension glaucoma. British Journal of Ophthalmology. 2008;92(9):1293-6.

332. Lin HJ, Tsai SC, Tsai FJ, et al. Association of interleukin 1β and receptor antagonist gene polymorphisms with primary open-angle glaucoma. Ophthalmologica Journal international d'ophtalmologie International journal of ophthalmology Zeitschrift fur Augenheilkunde. 2003;217(5):358-64.

333. Cuellar-Partida G, Springelkamp H, Lucas SE, et al. WNT10A exonic variant increases the risk of keratoconus by decreasing corneal thickness. Human molecular genetics. 2015;24(17):5060-8.

334. Chandra A, Abbas S, Raza S, et al. Polymorphism of CYP46A1 and PPARγ2 genes in risk prediction of primary open angle glaucoma among North Indian population. Middle East African Journal of Ophthalmology. 2016;23(2):172-6.

335. Macgregor S, Hewitt AW, Hysi PG, et al. Genome-wide association identifies ATOH7 as a major gene determining human optic disc size. Human molecular genetics. 2010;19(13):2716-24.

336. Wirtz MK, Keller KE. The role of the IL-20 subfamily in glaucoma. Mediators of Inflammation. 2016;2016.

337. Shiga Y, Akiyama M, Nishiguchi KM, et al. Genome-wide association study identifies seven novel susceptibility loci for primary open-angle glaucoma. Human molecular genetics. 2018;27(8):1486-96.

338. Aung T, Ocaka L, Ebenezer ND, et al. A major marker for normal tension glaucoma: association with polymorphisms in the OPA1 gene. Human Genetics. 2002;110(1):52-6.

339. Woo SJ, Kim DM, Kim JY, et al. Investigation of the association between OPA1 polymorphisms and normal-tension glaucoma in Korea. J Glaucoma. 2004;13(6):492-5.

340. Yao W, Jiao X, Hejtmancik JF, et al. Evaluation of the association between OPA1 polymorphisms and primary open-angle glaucoma in Barbados families. Molecular vision. 2006;12(73-75):649-54.

341. Yu-Wai-Man P, Stewart JD, Hudson G, et al. OPA1 increases the risk of normal but not high tension glaucoma. Journal of Medical Genetics. 2010;47(2):120-5.

342. Powell BL, Toomes C, Scott S, et al. Polymorphisms in OPA1 are associated with normal tension glaucoma. Molecular vision. 2003;9:460-4.

343. Mabuchi F, Tang S, Kashiwagi K, et al. The OPA1 Gene Polymorphism is Associated With Normal Tension and High Tension Glaucoma. American Journal of Ophthalmology. 2007;143(1):125-30.e2.

344. Ishikawa K, Funayama T, Ohtake Y, et al. Association between glaucoma and gene polymorphism of endothelin type A receptor. Molecular vision. 2005;11:431-7.

345. Porter LF, Urquhart JE, O'Donoghue E, et al. Identification of a novel locus for autosomal dominant primary open angle glaucoma on 4q35.1-q35.2. Invest Ophthalmol Vis Sci. 2011;52(11):7859-65.

346. Wang CY, Shen YC, Wei LC, et al. Polymorphism in the TNF-α(-863) locus associated with reduced risk of primary open angle glaucoma. Molecular vision. 2012;18:779-85.

347. Lin HJ, Tsai FJ, Chen WC, et al. Association of tumour necrosis factor alpha - 308 gene polymorphism with primary open-angle glaucoma in Chinese. Eye. 2003;17(1):31-4.

348. Mossböck G, Weger M, Moray M, et al. TNF-α promoter polymorphisms and primary open-angle glaucoma. Eye. 2006;20(9):1040-3.

349. Razeghinejad MR, Rahat F, Kamali-Sarvestani E. Association of TNFA -308 G/A and TNFRI +36 A/G gene polymorphisms with glaucoma. Ophthalmic research. 2009;42(3):118-24.

350. Bozkurt B, Mesci L, Irkec M, et al. Association of tumour necrosis factor-alpha -308 G/A polymorphism with primary open-angle glaucoma. Clinical and Experimental Ophthalmology. 2012;40(4):e156-e62.

351. Nowak A, Majsterek I, Przybyłowska-Sygut K, et al. Analysis of the expression and polymorphism of APOE, HSP, BDNF, and GRIN2B genes associated with the neurodegeneration process in the pathogenesis of primary open angle glaucoma. BioMed research international. 2015;2015.

352. Lin HJ, Tsai CH, Tsai FJ, et al. Transporter associated with antigen processing gene 1 codon 333 and codon 637 polymorphisms are associated with primary open-angle glaucoma. Molecular diagnosis : a journal devoted to the understanding of human disease through the clinical application of molecular biology. 2004;8(4):245-52.

353. Tsai FJ, Lin HJ, Chen WC, et al. A codon 31ser-arg polymorphism of the WAF-1/CIP-1/p21/tumour suppressor gene in Chinese primary open-angle glaucoma. Acta ophthalmologica Scandinavica. 2004;82(1):76-80.

354. Zhou Y, Shuai P, Li X, et al. Association of SOD2 polymorphisms with primary open angle glaucoma in a Chinese population. Ophthalmic Genet. 2015;36(1):43-9.

355. Abu-Amero KK, Kondkar AA, Mousa A, et al. Association of Mn-SOD mutation (c.47T > C) with various POAG clinical indices. Ophthalmic Genet. 2014;35(2):85-90.

356. The Blue Mountains Eye Study, Strange A, Bellenguez C, et al. Genome-wide association study of intraocular pressure identifies the GLCCI1/ICA1 region as a glaucoma susceptibility locus. Human molecular genetics. 2013;22(22):4653-60.

357. Liu H, Yang ZK, Li Y, et al. ABCB1 variants confer susceptibility to primary open-angle glaucoma and predict individual differences to latanoprost treatment. Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie. 2016;80:115-20.

358. Inagaki Y, Mashima Y, Funayama T, et al. Paraoxonase 1 gene polymorphisms influence clinical features of open-angle glaucoma. Graefe's Archive for Clinical and Experimental Ophthalmology. 2006;244(8):984-90.

359. Kim S, Kim K, Heo DW, et al. Expression-associated polymorphisms of CAV1-CAV2 affect intraocular pressure and high-tension glaucoma risk. Molecular vision. 2015;21:548-54.

360. Kuehn MH, Wang K, Roos B, et al. Chromosome 7q31 POAG locus: ocular expression of caveolins and lack of association with POAG in a US cohort. 2011.

361. Abu-Amero KK, Kondkar AA, Mousa A, et al. Lack of association of SNP rs4236601 near CAV1 and CAV2 with POAG in a Saudi cohort. 2012.

362. Emam WA, Zidan HE, Abdulhalim BEH, et al. Endothelial nitric oxide synthase polymorphisms and susceptibility to high-tension primary open-angle glaucoma in an Egyptian cohort. Molecular vision. 2014;20:804-11.

363. Tunny TJ, Richardson KA, Clark CV. Association study of the 5' flanking regions of endothelial-nitric oxide synthase and endothelin-1 gene in familial primary open-angle glaucoma. Clinical and Experimental Pharmacology and Physiology. 1998;25(1):26-9.

364. Micheal S, Ayub H, Islam F, et al. Variants in the ASB10 gene are associated with primary open angle glaucoma. PloS one. 2015;10(12).

365. Kimura Y, Akagi T, Miyake M, et al. Association between the CDKN2B-AS1 Gene and Primary Open Angle Glaucoma with High Myopia in Japanese Patients. Ophthalmic Genetics. 2016;37(2):242-4.

366. Chen Y, Lin Y, Vithana EN, et al. Common variants near ABCA1 and in PMM2 are associated with primary open-angle glaucoma. Nat Genet. 2014;46(10):1115-9.

367. Takano Y, Shi D, Shimizu A, et al. Association of Toll-like receptor 4 gene polymorphisms in Japanese subjects with primary open-angle, normal-tension, and exfoliation glaucoma. American journal of ophthalmology. 2012;154(5):825-32.e1.

368. Shibuya E, Meguro A, Ota M, et al. Association of toll-like receptor 4 gene polymorphisms with normal tension glaucoma. Investigative Ophthalmology and Visual Science. 2008;49(10):4453-7.

369. Park S, Jamshidi Y, Vaideanu D, et al. Genetic risk for primary open-angle glaucoma determined by LMX1B haplotypes. Investigative Ophthalmology and Visual Science. 2009;50(4):1522-30.

370. Vishal M, Sharma A, Kaurani L, et al. Genetic association and stress mediated down-regulation in trabecular meshwork implicates MPP7 as a novel candidate gene in primary open angle glaucoma. BMC Medical Genomics. 2016;9(1).

371. Chen J-H, Wang D, Huang C, et al. Interactive effects of ATOH7 and RFTN1 in association with adult-onset primary open-angle glaucoma. Investigative ophthalmology & visual science. 2012;53(2):779-85.

372. Tsai FJ, Lin HJ, Chen WC, et al. Insulin-Like Growth Factor-II Gene Polymorphism is Associated with Primary Open Angle Glaucoma. Journal of clinical laboratory analysis. 2003;17(6):259-63.

373. Gasten AC, Ramdas WD, Broer L, et al. A genetic epidemiologic study of candidate genes involved in the optic nerve head morphology. Investigative ophthalmology & visual science. 2012;53(3):1485-91.

374. Dimasi DP, Burdon KP, Hewitt AW, et al. Candidate gene study to investigate the genetic determinants of normal variation in central corneal thickness. Molecular vision. 2010;16:562-9.

375. Abu-Amero KK, Kondkar AA, Mousa A, et al. Analysis of catalase SNP rs1001179 in Saudi patients with primary open angle glaucoma. Ophthalmic Genet. 2013;34(4):223-8.

376. Majsterek I, Markiewicz L, Przybylowska K, et al. Association of MMP1-1607 1g/2g and TIMP1 372 T/C gene polymorphisms with risk of primary open angle glaucoma in a Polish population. Medical Science Monitor. 2011;17(7):CR417-CR21.

377. Sang J, Jia L, Zhao B, et al. Association of three single nucleotide polymorphisms at the SIX1-SIX6 locus with primary open angle glaucoma in the Chinese population. Science China Life sciences. 2016;59(7):694-9.

378. Fourgeux C, Martine L, Björkhem I, et al. Primary open-angle glaucoma: Association with cholesterol 24s-hydroxylase (CYP46a1) gene polymorphism and plasma 24-hydroxycholesterol levels. Investigative Ophthalmology and Visual Science. 2009;50(12):5712-7.

379. Wu M, Zhu XY, Ye J. Associations of polymorphisms of LOXL1 gene with primary open-angle glaucoma: A meta-analysis based on 5,293 subjects. Molecular vision. 2015;21:165-72.

380. Zanon-Moreno V, Zanon-Moreno L, Ortega-Azorin C, et al. Genetic polymorphism related to exfoliative glaucoma is also associated with primary open-angle glaucoma risk. Clinical and Experimental Ophthalmology. 2015;43(1):26-30.

381. Lin HJ, Tsai FJ, Hung P, et al. Association of E-cadherin gene 3'-UTR C/T polymorphism with primary open angle glaucoma. Ophthalmic research. 2006;38(1):44-8.

382. Funayama T, Mashima Y, Ohtake Y, et al. SNPs and interaction analyses of noelin 2, myocilin, and optineurin genes in Japanese patients with open-angle glaucoma. Invest Ophthalmol Vis Sci. 2006;47(12):5368-75.

383. Yousaf S, Khan MI, Micheal S, et al. XRCC1 and XPD DNA repair gene polymorphisms: A potential risk factor for glaucoma in the Pakistani population. Molecular vision. 2011;17:1153-63.

384. Vickers JC, Craig JE, Stankovich J, et al. The apolipoprotein epsilon4 gene is associated with elevated risk of normal tension glaucoma. Molecular vision. 2002;8:389-93.

385. Lake S, Liverani E, Desai M, et al. Normal tension glaucoma is not associated with the common apolipoprotein E gene polymorphisms. British journal of ophthalmology. 2004;88(4):491-3.

386. Ressiniotis T, Griffiths PG, Birch M, et al. The Role of Apolipoprotein E Gene Polymorphisms in Primary Open-angleGlaucoma. Archives of ophthalmology. 2004;122(2):258-61.

387. Fan BJ, Wang DY, Fan D, et al. SNPs and interaction analyses of myocilin, optineurin, and apolipoprotein E in primary open angle glaucoma patients. Molecular vision. 2005;11:625-31.

388. Mabuchi F, Tang S, Ando D, et al. The apolipoprotein E gene polymorphism is associated with open angle glaucoma in the Japanese population. Molecular vision. 2005;11:609-12.

389. Lam CY, Fan BJ, Wang DY, et al. Association of apolipoprotein E polymorphisms with normal tension glaucoma in a Chinese population. Journal of glaucoma. 2006;15(3):218-22.

390. Zetterberg M, Tasa G, Palmér MS, et al. Apolipoprotein E polymorphisms in patients with primary open-angle glaucoma. American journal of ophthalmology. 2007;143(6):1059-60.

391. Al-Dabbagh NM, Al-Dohayan N, Arfin M, et al. Apolipoprotein E polymorphisms and primary glaucoma in Saudis. 2009.

392. Saglar E, Yucel D, Bozkurt B, et al. Association of polymorphisms in APOE, p53, and p21 with primary open-angle glaucoma in Turkish patients. 2009.

393. Zanon-Moreno V, Asensio-Marquez EM, Ciancotti-Oliver L, et al. Effects of polymorphisms in vitamin E-, vitamin C-, and glutathione peroxidase-related genes on serum biomarkers and associations with glaucoma. Molecular vision. 2013;19:231-42.